Journal of Applied Statistics

ISSN: 0266-4763 (Print) 1360-0532 (Online) Journal homepage: http://www.tandfonline.com/loi/cjas20

Violent crime and incentives in the long-run: evidence from England and Wales

George Saridakis

To cite this article: George Saridakis (2011) Violent crime and incentives in the long-run: evidence from England and Wales, Journal of Applied Statistics, 38:4, 647-660, DOI: 10.1080/02664760903563619

To link to this article: https://doi.org/10.1080/02664760903563619

Published online: 10 Nov 2010.

Journal of Applied Statistics
Vol. 38, No. 4, April 2011, 647–660

Violent crime and incentives in the long-run: evidence from England and Wales

George Saridakis*

The Business School, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

(*Received 21 October 2008; final version received 11 November 2009*)

This study uses recent advances in time-series econometrics to investigate the non-stationarity and co-integration properties of violent crime series in England and Wales. In particular, we estimate the long-run impact of economic conditions, beer consumption and various deterrents on different categories of recorded violent crime. The results suggest that a long-run causal model exists for only minor crimes of violence, with beer consumption being a predominant factor.

Keywords: violent crime; time series; co-integration; weak exogeneity; error correction model

1. Introduction

Previous work analysing violent crime rates within the economic framework of crime [2,8] has mainly focused on the USA [20,32,33]. Using first-differenced data, this literature suggests that criminal justice policy and economic conditions have only a weak effect on violent crime. This casts doubt on the view that tougher penalties and a stronger economy will lead to less crimes of violence. The purpose of this paper is to re-examine this hypothesis using time-series data from England and Wales over the period 1960–2000. During this time, violent crime in England and Wales increased rapidly (Figure 1) reaching record levels in the 1990s, after which it became an increasingly important public policy issue. Indeed, the aggravated assault rate in England and Wales has been about 10% higher than the US rate since 1996 [10].

The present research adds to the literature by offering new and possibly more refined evidence from England and Wales, overcoming the limitations of earlier studies by Wolpin [36] and Field [11]: for a review, see Marris [23]. Since these earlier studies, newer tests have focused on the time-series properties of crime variables, considering whether the joint relationship between the variables is stationary, that is, whether the variables are co-integrated [4].[1] The focus on the non-stationarity and co-integration properties of various recorded violent crime series in England and Wales is the most innovative contribution of this paper. In particular, we make extensive use of the Vector Autoregressive (VAR) methodology to test for the existence of a relationship between

*Email: g.saridakis@lboro.ac.uk

ISSN 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/02664760903563619
http://www.informaworld.com

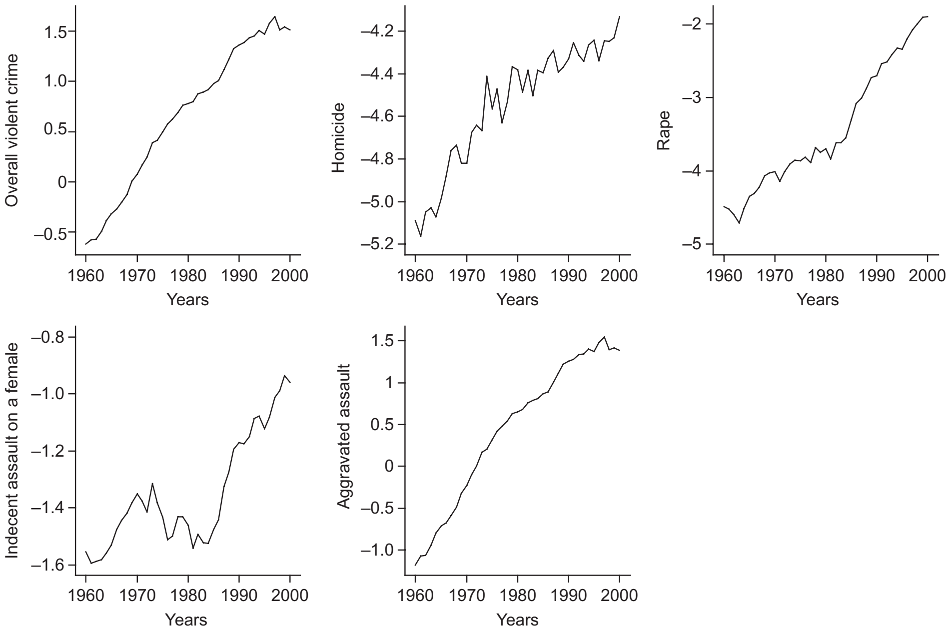

Figure 1. Log (type of violent crime per capita) versus time (years).
Note: Overall violent crime includes the offences of homicide, rape, indecent assault on a female and aggravated assault. Homicide comprises the recorded crimes of murder, manslaughter and infanticide. The offence of aggravated assault comprises the offences of felonious and malicious wounding. When aggregated, overall violent crime is largely made up of aggravated assault.

variables in levels [16,18]. In addition, the autoregressive distributed lag (ARDL) co-integration procedure developed by Pesaran *et al.* [27] is used as a supplementary approach. The value added of the latter technique is that it allows for testing of co-integration when the variables included in the crime model are a mixture of $I(1)$ and $I(0)$. This article demonstrates the usefulness of such methods to the study of criminal behaviour.

The violent crime equation draws upon the economics of crime literature including various deterrents (conviction rate and imprisonment rate) and proxies of legal-market opportunities (male unemployment rate and poverty rate). The economic model of crime predicts that an increase in the values of the deterrence variables should reduce criminal activity, whereas an increase in economic hardship should increase it. We also examine the effect of beer consumption on violent crime. Field [11] found a growth in beer consumption to be the most important factor in explaining the growth in crimes of violence against the person in England and Wales. Furthermore, the omission of the alcohol consumption variable from the crime equation may lead to biased estimates of the effect of unemployment on crime [32].

Our results indicate that a long-run causal model does not exist for the particularly serious crimes of violence, namely homicide, rape and indecent assault on a female. However, a long-run causal model does exist for aggravated assault. The normalized coefficients provide strong evidence for the deterrence hypothesis. Aggravated assault declines in response to increases in conviction and imprisonment rates. We also find that an increase in male unemployment and poverty rates generates an increase in aggravated assault. Of particular interest is the strong long-run effect of beer consumption on aggravated assault. *A priori*, one may expect such an effect to exist only in the short-run. Our interpretation is that habitual beer consumption may lead to long-run changes in behaviour, which can be associated with a rising level of violence or may be a proxy for social factors such as alienation and lack of opportunity.

The paper is organized as follows. Section 2 describes the data. Section 3 investigates the non-stationarity properties of the variables. Section 4 presents the statistical methodology. Section 5 discusses the empirical findings and the final section concludes the paper.

2. Data

We utilize annual time-series data for England and Wales over the period 1960–2000. These series were from various data sources. Violent crime and criminal justice recorded data were obtained from Criminal Statistics (CS). The British Crime Survey (BCS) provides an alternative source of data on violent crime and, in contrast to CS, shows a large fall in violent crime since 1995 [34]. Hale *et al.* [15, pp. 41–55] provide a comprehensive discussion of the two sources. We are aware that the CS data may be influenced by the propensity of victims to report the incident and/or variations in recording practices among police forces. However, the data provided by the BCS are too short to allow for a meaningful examination of the time-series properties of the data set. The rest of the data used in this study has been collected from the following sources: Population Trends provide data for the national population; the male unemployment rate has been constructed from the Annual Abstract of Statistics; the Institute for Fiscal Studies provides information on poverty rate and finally, data on the amount of beer released to the public (i.e. restaurants, pubs, shops, etc.) have been obtained from the Annual Abstract of Statistics. Table 1 provides details on how the variables to be used in the regressions are constructed.

3. Unit root analysis

In performing co-integration analysis, the first step is to determine the order of integration of the individual variables. The decision about the order of integration (i.e. the number of times a variable needs to be differenced in order to make it a stationary series) is crucial to the choice of technique used to estimate the violent crime model. An incorrect decision will lead to the estimation of a misspecified model and hence misleading conclusions. We start by investigating the unit roots in the violent crime series, such that the order of integration can be determined. Previous research tends to give conflicting results regarding the order of integration of the crime series. Pyle and

Table 1. Definitions of the variables.

$V_{j,t}$	Violent crime *per capita* in England and Wales
$C_{j,t}$	Conviction rate: The number of convictions for violent crime in England and Wales divided by the number of recorded violent crime
$I_{j,t}$	Imprisonment rate: The number imprisoned for violent crime in England and Wales divided by the number convicted for violent crime
U_t	UK male unemployment rate
P_t	UK poverty rate: The proportion of households below 60% median income
B_t	UK beer consumed by the total persons aged 16–74: Hectolitres of beer released to the public divided by the total persons aged 16–74

Notes: All the variables are measured in natural logarithms. V_j, C_j and I_j stand for jth violent crime ($j = 1$, overall violent crime; $j = 2$, homicide; $j = 3$, rape; $j = 4$, indecent assault on a female and $j = 5$, aggravated assault). The correlation matrix and the variance inflation factors indicate that multicollinearity does not cause problems in the estimations. A demographic variable was also constructed to express the proportion of males aged 15–24. When the variable was tested for the order of integration, we found that the variable was $I(2)$. The differenced demographic variables could have been used in our specification, but would have lacked theoretical justification, that is, the level of violent crime may depend on the proportion of young men in the population, but not on their first difference. For this reason, we decided not to include this variable in our model.

Deadman [30] found that the property crime data were $I(2)$. If that is correct, then crime rates would grow indefinitely, even in a stationary long-run. However, other scholars [14,29] showed that the crime data were in fact $I(1)$. Wolpin [36] and Field [11] did not determine the existence and nature of non-stationarity in the crime data.

3.1 *The ADF and PP tests*

In this paper, the augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests were applied to test the order of integration of the variables [7,28]. The ADF and PP tests statistics for the levels and first differences of the variables, computed over the period 1963–2000, are reported in Table 2. The critical values for both tests are identical and are provided by MacKinnon [22]. Our results suggest that the violent crime series are non-stationary in the level form but stationary after first differencing. However, both tests suggest that the conviction rate and imprisonment rate for homicide are stationary in levels. All the other explanatory variables considered in this study were found to be $I(1)$.

3.2 *The Perron test*

There is one point, however, that deserves further attention; there appears to be a break in the rape and indecent assault on a female series in 1980 which leads to permanently higher mean level of these processes. As far as we can ascertain, the break is not due to a change in how the relevant variables are defined, but rather due to a change in the law (Sexual Offences Amendment Act in 1976), more general administrative improvements and changes to the police's approach

Table 2. Tests for stationarity.

	ADF		PP	
Variable	Levels	First difference	Levels	First difference
$V_{1,t}$	0.336	−4.711	0.253	−6.745
$V_{2,t}$	−2.523	−6.335	−3.289	−9.336
$V_{3,t}$	−1.483	−4.381	−1.593	−6.898
$V_{4,t}$	−1.343	−3.891	−1.188	−5.136
$V_{5,t}$	0.208	−4.936	0.118	−6.819
$C_{1,t}$	−1.691	−4.067	−1.425	−4.605
$C_{2,t}$	−4.603	−8.206	−5.055	−8.386
$C_{3,t}$	−1.327	−6.209	−1.546	−8.268
$C_{4,t}$	−1.043	−4.432	−0.916	−5.389
$C_{5,t}$	−2.072	−4.403	−1.739	−4.673
$I_{1,t}$	−1.338	−4.879	−1.404	−6.289
$I_{2,t}$	−4.182	−6.416	−4.009	−6.725
$I_{3,t}$	−3.094	−5.08	−3.015	−6.177
$I_{4,t}$	−2.335	−5.151	−2.31	−5.562
$I_{5,t}$	−1.282	−4.742	−1.403	−6.612
U_t	0.106	−5.931	−0.882	−4.560
P_t	−3.164	−3.119	−1.718	−2.609
B_t	−1.369	−4.139	−1.309	−5.507

Notes: The figures reported are the t-ratios of the estimated coefficients (sample period 1963–2000). One lag of the dependent variable was used for conducting the ADF test statistic, with the exception of variable U_t where two lags of the dependent variable were required to clean the residuals in the ADF regression model. A time trend was included when conducting both the ADF and the PP tests, since all the series are clearly trended. The MacKinnon critical values are −3.533 (5%) and −3.198 (10%).

to sexual offences around 1980. An increase in the number of rape cases reported and charged is expected in response to the above changes. Perron [26] suggested that in such cases where a break exists, the Dickey–Fuller test is biased towards non-rejection of the non-stationary null hypothesis. In light of this, we carry out the Perron test [26] for a unit root in the presence of a structural break; this tests the null hypothesis of a non-stationary series subject to a single-pulse intervention at a known date against the alternative of a stationary process subject to a shift.

The test is performed using a two-stage procedure: see Charemza and Deadman [6, pp. 115–122] for a exposition of the procedure. Briefly, in the first stage, the ordinary least squares (OLS) residuals from the regression under the alternative hypothesis are computed. If the alternative hypothesis of a stationary process is true, then this regression should yield residuals that have been shorn of the influence of the intervention. In the second stage, differences of these residuals are regressed on the lagged residuals and a pulse variable. Here, the t-values on the lagged residuals take the values of -0.86 and -0.76 from the second stage regression for rape and indecent assault on a female, respectively. For a test at the 5% significance level with 41 observations, the lower and upper bounds' critical values are -3.56 and -3.51 [6, Table 6, p. 303]. The calculated t-values lie above the upper level critical value for the test, such that the null hypothesis of the series being non-stationary cannot be rejected, which is consistent with the ADF and PP tests (results are available upon request).

4. Estimation procedure

Using classical estimation methods, such as OLS, to estimate relationships among variables with unit roots can give misleading inferences and lead to spurious regression. However, one can difference all the $I(1)$ variables to make them $I(0)$ and then model the relationship between the differenced variables [3]. The drawback of this procedure, however, is that information about the long-run is lost.[2] The study by Field [11], for example, explores relationships between growth rates rather than between levels. Recent advances in the time-series analysis deal with estimation and inferences in the presence of $I(1)$ variables. Johansen [18,19] describes a full information maximum-likelihood procedure based on the principle that variables that share the same stochastic trend can be combined linearly in a way that eliminates the stochastic trend. This linear combination can be interpreted as the long-run link between the non-stationary series.

The procedure to estimate the co-integrating vector(s) is derived from a VAR model in levels, which is represented by Equation (1):

$$z_t = \sum_{i=1}^{p} \Phi_i z_{t-i} + \Psi D_t + u_t, \quad t = 1, \ldots, T, \tag{1}$$

where z_t is the vector of the stochastic variables $(V_{j,t}, C_{j,t}, I_{j,t}, U_t, P_t, B_t)$; Φ_i a 6×6 matrix of coefficients on the ith lag of z_t; p the maximal lag length; Ψ a $6 \times s$ matrix of coefficients on D_t, where D_t is a vector of s deterministic variables (such as a constant term and a trend) and u_t is a vector of error terms each of which is assumed to be normally, independently and identically distributed. Throughout, z is restricted to be integrated of order one, denoted $I(1)$. In the presence of co-integration, Equation (1) can be written as a vector error correction model, that is,

$$\Delta z_t = \pi z_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + \Psi D_t + u_t, \tag{2}$$

where π and Γ_i are given by

$$\pi = \left(\sum_{i=1}^{p} \Phi_i \right) - I_6, \tag{3}$$

$$\Gamma_i = -(\Phi_{i+1} + \cdots + \Phi_p), \quad i = 1, \ldots, p-1, \tag{4}$$

where I_6 is the identity matrix of dimension six and Δ the difference operator. However, given the fact that there is no explicit theoretical model proposed, and without satisfactory equations for all the endogenous variables, an estimation of the full system is not appropriate. In order to reduce the dimension of the system and improve efficiency, we determine the co-integrating rank by an analysis of a partial system [16]. Based on the existing literature, we suspect that there may be a possible simultaneity between crime and deterrence. Here, the main point is that if the criminal justice system (CJS) resources remain constant, an increase in criminal activity may overburden existing resources resulting in a reduction in the effectiveness of the CJS in producing sanctions [21].

Using the Wu–Hausman statistic [17,37], we found that only the conviction rate, $C_{j,t}$, is endogenous in each of the violent crime equations with a sole exception in the homicide equation where both deterrent variables were found to be exogenous (results are available upon request): lagged values of the crime and deterrent variables were used as instruments for this test. Our modelling strategy, hence, involves partitioning $z_t = (y_t', x_t')'$, where $y_t = (V_{j,t}, C_{j,t})'$ is treated as an $I(1)$ vector of endogenous variables and $x_t = (I_{j,t}, U_{j,t}, P_t, B_t)'$ as a vector of weakly exogenous $I(1)$ variables, in that there are no feedbacks from the long-run disturbances to these variables. Then, the conditional model and the marginal model can be written as follows:

$$\Delta y_t = b_{y0} + b_{y1}t + \omega \Delta x_t + \sum_{i=1}^{p-1} \tilde{\Gamma}_{yi} \Delta z_{t-i} + \pi_y z_{t-1} + \tilde{u}_{yt}, \tag{5}$$

$$\Delta x_t = b_{x0} + \sum_{i=1}^{p-1} \Gamma_{xi} \Delta z_{t-i} + u_{xt}, \tag{6}$$

where Equation (5) contains an unrestricted intercept ($b_{y0} \neq 0$) and a restricted trend ($b_{y1} = -\pi_y \gamma$) to rule out a quadratic trend [19, pp. 156–163]. For any specified reduced rank $r (0 \leq r \leq 2)$, the matrix π_y is of potentially reduced rank r, and may be written as $a_y \beta'$ where the $(2, r)$ loading matrix a_y and the $(6, r)$ matrix of co-integrating vectors β are each full-column rank and identified up to an arbitrary (r, r) non-singular matrix. Harbo *et al.* [16] provides tabulated critical values for the λ_{trace} statistics, which tests the null hypothesis that $r \leq q$, for some q, against a general alternative. Finally, Equation (6) allows for feedbacks from Δy to Δx, but does not allow for level feedbacks.

5. Estimation results

5.1 *Model specification*

We use the above methodology to study the overall violent crime and the three individual crimes of violence: rape, indecent assault on a female and aggravated assault. This statistical approach is not appropriate to study homicide since two of its major determinants $(C_{2,t}, I_{2,t})$ are found to be $I(0)$, and so the procedure developed in Pesaran *et al.* [27] is used. The first task of our modelling sequence is to specify the relevant order of lag(p) for the underlying VARs in these variables. Estimation of an unrestricted VAR(3) in levels, of y_t, a set of exogenous variables x_t and including an intercept, indicates that a lag length of one ($p = 1$) is appropriate for each jth violent crime model. This is supported by the Schwarz–Bayesian criterion and the adjusted-likelihood ratio test.

Table 3 presents the residual diagnostic tests for the estimated unrestricted VARs. Since we assumed uncorrelated errors, there should be no serial correlation in the residuals. Table 3 supports this assumption. Furthermore, we assumed that the errors are identically distributed, such that there is no heteroscedasticity. Table 3 also shows that the ARCH tests of order one support the null hypothesis of no ARCH effects. A final test presented in Table 3 concerns the normality assumption. The Jarque–Bera test indicates that non-normality is exclusively in the equations for $C_{1,t}$, $C_{4,t}$ and $C_{5,t}$ due to excess kurtosis. Since co-integration results have been found to be quite robust to the excess kurtosis [12], we regard the present model specification to be acceptable. Although not presented here, the inverse roots of the AR(1) models indicate that the roots lie inside in the unit circle and that the estimated VAR models satisfy the stability condition.

5.2 *Determination of the co-integration rank*

The values of the likelihood-ratio tests for $r = 0$ based on the partial models together with their associated 95% critical values are reported in Table 4. Under the null hypothesis of no co-integration ($r = 0$) each system has $p - r = 6$ common trends, of which two ($p_1 - r = 2$) are in the partial system. From Table 2 of Harbo *et al.* [16] with $p_2 = 4$ and $p_1 - r = 2$, the 95% quantile is 44.5. The λ_{trace} statistic rejects the null hypothesis of no co-integrating vectors in favour of exactly one co-integrating vector for overall violent crime and aggravated assault (it should be mentioned that overall violent crime is largely made up by aggravated assault). However, the null hypothesis of $r = 0$ is not rejected for rape and indecent assault on a female Thus, for the last two violent crime models π_y in Equation (5) is completely deficient in rank, rank(π_y) = 0. This finding is in line with recent research on US violent crime, reported in [33].

We also examine the degree to which the co-integration results for the sexual offences are robust when a step dummy that is zero up to 1979 and unity thereafter enters the models unrestrictedly, that is, outside the co-integration space. The accumulated value of this dummy is a broken linear trend and that will influence the limit distribution of the trace statistic: this is discussed in [18, p. 84]. To surmount this, Rahbek and Mosconi [31] suggest including the broken linear trend in

Table 3. Residual diagnostic tests for the VAR models.

Endogenous variable	LMSC(2)	ARCH(1)[a]	Skewness[b]	Kurtosis[b]	$N(2)$
$V_{1,t}$	2.08[0.1]	0.19[0.7]	0.04	2.93	0.17[0.9]
$C_{1,t}$	1.99[0.1]	0.27[0.6]	0.78	5.41	13.76[0.0]*
$V_{3,t}$	1.35[0.3]	0.01[0.9]	−0.15	2.39	0.75[0.7]
$C_{3,t}$	1.44[0.3]	1.03[0.3]	0.52	3.22	1.19[0.4]
$V_{4,t}$	0.61[0.5]	0.05[0.8]	0.10	2.12	1.36[0.5]
$C_{4,t}$	0.93[0.4]	0.04[0.8]	1.26	8.05	53.2[0.0]*
$V_{5,t}$	1.84[0.2]	0.01[0.9]	−0.17	2.58	0.49[0.8]
$C_{5,t}$	2.43[0.3]	0.23[0.6]	0.40	5.20	9.15[0.0]*

Notes: LMSC(2) is a test for up to second-order serial correlation $F(2,31)$; ARCH(1) is a test for conditional heteroscedasticity [9] $F(1,32)$; $N(2)$ is the Jarque–Bera test for normality and is asymptotically distributed as $\chi^2(2)$.

[a]We also examined the individual series by the ARCH test for residual conditional heteroscedasticity of orders two and three. The ARCH tests of orders two and three support the null hypothesis of no ARCH effects.

[b]The skewness and kurtosis of the normal distribution are zero and three, respectively.

*In square brackets are the P-values.

Table 4. Trace test of the co-integration rank.

$(V_{1,t}, C_{1,t})$ conditional on $(I_{1,t}, U_t, P_t, B_t)$		
Hypothesis	$r = 0$	$r \leq 1$
Eigenvalues	0.580	0.367
Trace statistic	53*	18.3
95% Critical value	44.5	22.9
$(V_{3,t}, C_{3,t})$ conditional on $(I_{3,t}, U_t, P_t, B_t)$		
Hypothesis	$r = 0$	$r \leq 1$
Eigenvalues	0.399	0.338
Trace statistic	36.9	16.5
95% Critical value	44.5	22.9
$(V_{4,t}, C_{4,t})$ conditional on $(I_{4,t}, U_t, P_t, B_t)$		
Hypothesis	$r = 0$	$r \leq 1$
Eigenvalues	0.361	0.186
Trace statistic	26.2	8.25
95% Critical value	44.5	22.9
$(V_{5,t}, C_{5,t})$ conditional on $(I_{5,t}, U_t, P_t, B_t)$		
Hypothesis	$r = 0$	$r \leq 1$
Eigenvalues	0.530	0.381
Trace statistic	49.4*	19.2
95% Critical value	44.5	22.9

Notes: The underlying VAR models are of order one and contain unrestricted intercept and restricted trend with $(V_{j,t}, C_{j,t})$ treated as endogenous $I(1)$ variables and $(I_{j,t}, U_t, P_t, B_t)$ treated as weakly exogenous $I(1)$ variables. The statistics are computed using 40 observations for the period 1961–2000. The critical values for the trace statistics are taken from [16, Table 2].
*Significant at the 5 % level.

the model in levels and restrict it to the co-integration space, that is,

$$\Delta y_t = \omega \Delta x_t + \sum_{i=1}^{p-1} \tilde{\Gamma}_{yi} \Delta z_{t-i} + \pi_y^* \begin{bmatrix} z_{t-1} \\ t \\ \tilde{D}_t \end{bmatrix} + \Psi_y D_t + \tilde{u}_{yt}, \tag{7}$$

where

$$\tilde{D}_t = \sum_{i=1}^{t} D_i = \begin{cases} 0, & t \leq 1979, \\ t - 1979, & t = 1980, \ldots, 2000, \end{cases}$$

and D_t is a 2×1 vector containing a constant and a shift dummy variable ($d_t = 1_{\{t \geq 1980\}}$). In this case, the null hypothesis of no co-integration is rejected, whereas the hypothesis that $r \leq 1$ is not rejected (results available upon request). However, the maximum-likelihood estimate of the co-integrating vectors (normalized) for both models show that the variables are statistically insignificant or/and bear signs contrary to the theory.[3]

5.3 *Parameter estimates for the partial models*

The resultant error correction coefficient (also known as the loading coefficient) α_1 and co-integrating parameter β_e are reported in Table 5. A number of pertinent findings are apparent. First, the estimated elasticities of violent crime with respect to conviction rate and imprisonment rate are found to be statistically significant and in the neighbourhood of -0.3; their statistical significance contrasts with the findings of Wolpin [36].[4] Second, the results also indicate that

Table 5. Estimates for the partial models.

	$V_{1,t}$	$C_{1,t}$	$I_{1,t}$	U_t	P_t	B_t	Trend
β_e'	1	0.26[4.43]	0.25[9.23]	−0.16[16.89]	−0.24[10.31]	−1.12[11.08]	−0.04[14.31]
α_1'	−0.59(0.11)	0.68(0.21)					
Estimates of α_2 from the marginal model							
α_2'			−1.03(1.19)	−0.3(2.11)	−0.71(0.34)	0.20(0.37)	
	$V_{5,t}$	$C_{5,t}$	$I_{5,t}$	U_t	P_t	B_t	Trend
β_e'	1	0.30[3.16]	0.35[13.65]	−0.20[13.95]	−0.13[2.18]	−1.40[7.59]	−0.05[7.28]
α_1'	−0.50(0.10)	0.56(0.18)					
Estimates of α_2 from the marginal model							
α_2'			−1.32(0.70)	0.69(1.15)	−0.38(0.21)	0.15(0.18)	

Notes: Long-run estimates are normalized values. The LR statistics are given in square brackets, distributed as $\chi^2(1)$; all the estimates are statistically significant at the 5% level with the exception of the conviction rate and poverty rate coefficients that were found to be significant at or around the 10% level in the aggravated assault equation. The error correction terms have associated standard errors attached in parentheses; it is found that the null hypothesis that $\alpha_2 = 0$ cannot be rejected at the 5% level.

an increase in male unemployment and poverty serves to raise criminal activity. Some empirical studies, however, have found that unemployment exerts an insignificant (or even negative) effect on violent crime. For example, the US studies by Levitt [20] using state-level panel data and Mocan and Rees [24] using cross-section data found no significant impact of the unemployment rate on violent crime. The latter research, however, showed that poverty is positively related to the likelihood of committing assaults. But a third noteworthy finding is that there is a substantial positive long-run relationship between beer consumption and aggravated assault. This finding contrasts with Ormerod *et al.*'s [25] hypothesis that there is no apparent association between beer consumption and violent crime. Their argument was based on the observation that beer consumption has steadily declined over time while violent crime has increased, but the alcohol–crime relationship was not tested empirically.[5] We suggest that the decline in beer consumption during recent years may indicate a shift in cultural characteristics, and hence it may require some time to pass before the effects show up on violent crime rates. Finally, the error correction terms indicate fairly fast error-correction behaviour. For example, the estimated coefficient value of −0.59 suggests that 59% of any deviation from the long-run relationship is eliminated each year.

The misspecification test for weak exogeneity of $I_{j,t}$, U_t, P_t and B_t can be tested in the marginal model, Equation (6), by regression of the changes of these variables on $\beta_e' z_{et}$, the lagged differences of all variables and a constant. This is illustrated in [16]. The marginal model includes a substantial number of lags to address problems with autocorrelation. The least-squares estimates of α_2 appear in Table 5, because of the large standard errors, the null hypothesis that the associated α_2 coefficient is different from zero cannot be rejected. Therefore, it is reasonable to assume that the imprisonment rate, male unemployment, poverty and beer consumption are weakly exogenous in the overall violent crime and aggravated assault models.

Table 6 reports the overall violent crime and aggravated assault OLS estimates for the error correction models (ECMs). Of the economic variables, only male unemployment is found to have a significant and positive short-run effect, but the effect is fairly small numerically.[6] This suggests that violent crime does not adjust instantly to short-run economic conditions, which seems reasonable if time is required to exhaust short-term financial resources [5] before engaging

Table 6. OLS estimates of the ECMs.

Variable	$\Delta V_{1,t}$	$\Delta V_{5,t}$
ΔI_t	−0.21*(0.05)	−0.25*(0.05)
ΔU_t	0.05*(0.02)	0.06*(0.03)
ΔP_t	−0.02(0.10)	−0.08(0.12)
ΔB_t	1.01*(0.17)	1.22*(0.20)
z_{t-1}	−0.59*(0.11)	−0.50*(0.10)
Constant	−1.31*(0.24)	−1.32*
R^2	0.71	0.71
σ	0.03	0.03
N(1961–2000)	40	40
LMSC(2,32)	0.55[0.58]	1.60[0.22]
N(2)	0.46[0.80]	0.85[0.65]
H(1,38)	0.00[0.97]	0.1E−4[0.99]
RR(1,33)	3.04**[0.09]	0.76[0.40]

Notes: Standard errors of short-run coefficients are in parentheses: P-values are in square brackets; σ is the standard error of the regression; LMSC, N, H and RR are, respectively, tests of serial correlation, normality of residuals, heteroscedasticity and functional form. All are distributed as F-variates, with degrees of freedom given in parentheses, except N, which is distributed as $\chi^2(2)$ variate. The error correction term is: $z = V_{-1} + \beta_1 C_{-1} + \beta_2 I_{-1} + \beta_3 U_{-1} + \beta_4 P_{-1} + \beta_5 B_{-1} + \beta_6$Trend.
*Statistically significant at the 5% level.
**Statistically significant at the 10% level.

in criminal activity. Our results are similar to the short-run estimates of [11] with respect to beer consumption. Finally, the CUSUM and CUSUMSQ tests suggest that the parameters are stable (results available upon request).

5.4 *Is there a long-run model for homicide?*

Based on the earlier discussion, a different procedure should be adopted to test for the existence of a relationship between variables in levels when the variables included in a model are a mixture of $I(1)$ and $I(0)$. Therefore, the bounds testing procedure to co-integration developed by Pesaran *et al.* [27] is used to study homicide. This is described analytically in Appendix 1. In testing the null of no co-integration, we must first decide the order of lags on the first-differenced variables in Equation (A1). Owing to the size of the sample, we impose one and two lags on the first differences of each variable and compute the F-statistic for the joint significance of lagged levels of the variables. The computed F-statistic for each order of lag is reported in Table 7 along with the critical values at the bottom of the table. The test outcomes do not vary with the choice of the lag order and since the F-statistic lies below the 0.05 lower bound, no evidence of a long-run relationship is found.

We also consider an alternative specification where changes in the male unemployment rate are used rather than the level of the male unemployment rate [5,13]. This substitution does not affect our conclusions and therefore the results based on the latter variable are not reported.

6. Conclusions

In this paper, we use recent developments in time-series co-integration analysis to estimate the long-run impact of economic conditions, various deterrents and beer consumption on different categories of violent crime in England and Wales. The focus on the long-run relations makes this paper different from the majority of the economic literature on violent crime which focuses on the US and uses first-differenced data, that is, year-to-year growth rates of crime [20]. Our paper

Table 7. F-statistic for testing the existence of a long-run in homicide equation.

	Unrestricted intercept and no trend	Unrestricted intercept and unrestricted trend
Order of lag	F-statistics	F-statistics
1	0.515*	1.935*
2	2.248*	0.739*

Notes: The F-statistic is used to test for the joint significance of the coefficients of the lagged levels in the ARDL-ECM. The relevant critical value bounds are given in [27, Table CI.iii and CI.v]. The critical values are 2.62–3.79 (without trend) and 3.12–4.25 (with unrestricted trend) at the 5% significance level.
* F-statistic lies below the 0.05 lower bound.

overcomes limitations of earlier studies by Wolpin [36] and Field [11] and provides new and possibly more refined evidence from England and Wales.

Our results suggest that there are no long-run models for serious crimes of violence. Thus, we argue that the ‘rational choice’ economic model of crime, which has been widely used for modelling property crimes (e.g. burglary, car theft, etc.), may not be an appropriate framework for explaining and subsequently forecasting serious violent behaviour in the long-run. Moreover, providing doubts on the concept of a long-run equilibrium for serious violent crime through our analysis, minimizes substantially the importance of using ECM for modelling these offences. From a policy perspective, costly deterrence-based policies adopted by governments, for instance, are unlikely to alter serious violent crime in the long-run. However, a long-run model does exist for aggravated assault. We found that aggravated assault declines in response to increases in conviction and imprisonment rates, a result that supports the deterrence hypothesis. On the other hand, we found that increases in male unemployment rates and poverty rates generated an increase in aggravated assault. The main result however, is the significant long-run impact of beer consumption on aggravated assault. In light of finding this long-run relationship between beer consumption and aggravated assault, we suggest that efforts targeted at beer consumption may be viewed as an effective method of crime prevention. However, due to the close association of drinking with cultural and social characteristics, the impact on violent crime rates of any intervention may require some time to take effect. Turning to short-run estimates, we found that aggravated assault is not caused by the short-run economic conditions, but beer consumption was found to be the most important factor in explaining growth in aggravated assault.

Some limitations and possible extensions of the current study should be noted. First, CS data used in this paper may be influenced by the propensity of victims to report the incident and/or variations in recording practices among police forces. Crime victimization surveys such as the BCS may give a more accurate estimate than official statistics, but the data provided by the BCS are too short to allow for a meaningful examination of the time-series properties of the data set. Second, an important feature of the time-series pattern of crime is the nature of the seasonality in the data. However, a study of this issue requires quarterly data for crime, deterrence and socio-economic variables. Unfortunately, similar to the overwhelming preponderance of prior published research in the area, we only have annual data and so cannot effectively address this issue. Third, the use of regional crime data and recent developments in panel co-integration [1] may be fruitful. For example, the range of covariates in the time-series analysis is restricted due to severe limitations in the available degrees of freedom. Furthermore, some variables may exhibit significant variability at a local level, whereas national data average across all of these localized fluctuations. However, data limitations may also frustrate this venture. Fourth, co-integration analysis for partial models in the presence of structural breaks, as well as non-linear time-series analysis (e.g. there may be threshold effects whereby the long-run relationships change if certain variables are above given thresholds levels) could be of interest; however, statistical theory in this area is still in its infancy. Fifth, this paper examines the relationship between violent crime and beer consumption. But there are other alcoholic beverages (e.g. spirits and wine) that may also be associated with violent crime. This is an issue for future research. Finally, the empirical model of violent crime could be usefully extended to other countries in order to establish whether the results obtained here are unique to England and Wales or have wider applicability elsewhere.

Acknowledgements

We would like to thank for their helpful comments Wojciech Charemza, Mike Clements, Derek Deadman, Baibing Li, Robin Marris, Steve Pudney, Kavita Sirichand, Mark Tippett and Paul Tracey. In addition, the paper benefited from comments from the participants of seminars and conferences at the University of Leicester, the Institute for Social and Economic Research and the 4th European Society of Criminology Conference. The usual disclaimer applies.

Notes

1. Co-integration methods have been used in studying crime in the USA [35], as well as property crime in the UK [14].
2. When the analysis is conducted in levels, the results show that there is a long-run relationship among the variables, whereas when the analysis is conducted in differences the results measure the effect of short-term changes (information about the trend in the level of the series is discarded).
3. The sole exceptions are the estimated imprisonment rate (-0.097) and poverty rate (0.323) elasticity parameters in the indecent assault on a female model were found to be statistically significant and in line with the theory.
4. Wolpin used national crime statistics for E&W for the period 1894–1967 and estimated a single-equation model to assess the effect of deterrence and economic conditions on specific types of violent crime. Although Wolpin's study provides an important contribution to the economics of crime literature, it generally lacks statistical testing for reliability of the regression models and the inference which were derived from them.
5. Ormerod *et al.* [25] provide, however, an innovative methodological approach to the process by which crime spreads (or contracts) over time, and makes an interesting theoretical contribution to the role of social interaction between agents.
6. Field [11] also found a small positive and significant relationship between unemployment and violence against the person.

References

[1] B. Baltagi, T. Fomby, and R.C. Hill, *Nonstationary Panels, Cointegration in Panels and Dynamic Panels*, Advances in Econometrics, Elsevier, Amsterdam, 2000.
[2] G. Becker, *Crime and punishment: An economic approach*, J. Polit. Econ. 76(2) (1968), pp. 169–217.
[3] G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.
[4] C. Britt, *Testing theory and the analysis of time series data*, J. Quant. Criminol. 17(4) (2001), pp. 343–357.
[5] D. Cantor and K. Land, *Unemployment and crime rates in the post-world war II United States: A theoretical and empirical analysis*, Am. Sociol. Rev. 50(3) (1985), pp. 317–332.
[6] W. Charemza and D. Deadman, *New Directions in Econometric Practice*, 2nd ed., Edward Elgar, Cheltenham, 1997.
[7] D.A. Dickey and W.A. Fuller, *Distributions of the estimators for autoregressive time series with a unit root*, J. Am. Stat. Assoc. 74 (1979), pp. 427–431.
[8] I. Ehrlich, *Participation in illegitimate activities: A theoretical and empirical investigation*, J. Polit. Econ. 81(3) (1973), pp. 521–565.
[9] R.F. Engle, *Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation*, Econometrica 50 (1982), pp. 987–1008.
[10] D.P. Farrington, P. Langan, and M. Tonry, *Cross-National Studies in Crime and Justice*, US Department of Justice, Bureau of Justice Statistics, 2004.
[11] S. Field, *Trends in Crime and Their Interpretation: A Study of Recorded Crime in Post-War England and Wales*, Home Office Research Study Vol. 119, Home Office, London, 1990.
[12] J. Gonzalo, *Five alternative methods of estimating long-run equilibrium relationships*, J. Econ. 60 (1994), pp. 203–233.
[13] D. Greenberg, *Time series analysis of crime rates*, J. Quant. Criminol. 17(4) (2001), pp. 291–327.

[14] C. Hale, *Crime and business cycle in post-war Britain revisited*, Br. J. Criminol. 38(4) (1998), pp. 681–698.
[15] C. Hale, K. Hayward, A. Wahidin, and E. Wincup, *Criminology*, Oxford University Press, Oxford, 2005.
[16] I. Harbo, S. Johansen, B. Neilsen, and A. Rehbek, *Asymptotic inference on cointegrating rank in partial systems*, J. Bus. Econ. Stat. 16 (1998), pp. 388–399.
[17] J.A. Hausman, *Specification test in econometrics*, Econometrica 46 (1978), pp. 1251–1272.
[18] S. Johansen, *Statistical analysis of cointegration vectors*, J. Econ. Dyn. Control 12 (1988), pp. 231–254.
[19] S. Johansen, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford, 1995.
[20] S. Levitt, *The effect of prison population size on crime rates: Evidence from prison overcrowding litigation*, Quart. J. Econ. 111(2) (1996), pp. 319–351.
[21] S. Machin and C. Meghir, *Crime and economic incentives*, Working Paper, The Institute for Fiscal Studies, 2000.
[22] J.G. MacKinnon, *Numerical distribution functions for unit root and cointegration tests*, J. Appl. Econ. 11(6) (1996), pp. 601–618.
[23] R. Marris, *Survey of the Literature on the Criminological and Economic Factors Influencing Crime Trends*, Modeling Crime and Offending: Recent Developments in England and Wales, Home Office Occasional Paper 80, Home Office, London, 2003.
[24] H.N. Mocan and D.I. Rees, *Economic conditions, deterrence and juvenile crime: Evidence from micro data*, Am. Law Econ. Rev. 7(2) (2005), pp. 319–349.
[25] P. Ormerod, C. Mounfield, and L. Smith, *Non-Linear Modelling of Burglary and Violent Crime in the UK*, Modelling Crime and Offending: Recent Developments in England and Wales, Home Office Occasional Paper 80, Home Office, London, 2003.
[26] P. Perron, *Testing for a unit root in a time-series with a changing mean*, J. Bus. Econ. Stat. 8 (1990), pp. 153–162.
[27] H. Pesaran, Y. Shin, and R. Smith, *Bounds testing approaches to the analysis of level relationships*, J. Appl. Econ. 16 (2001), pp. 289–326.
[28] P.C.B. Phillips and P. Perron, *Testing for a unit root in time series regression*, Biometrika 75 (1988), pp. 335–346.
[29] S. Pudney, D. Deadman, and D. Pyle, *The relationship between crime, punishment and economic conditions: Is reliable inference possible when crimes are under-recorded?*, J. R. Stat. Soc. Ser. A Stat. Soc. 163(1) (2000), pp. 81–97.
[30] D. Pyle and D. Deadman, *Crime and business cycle in post-war Britain*, Br. J. Criminol. 34 (1994), pp. 339–357.
[31] A. Rahbek and R. Mosconi, *Cointegration rank inference with stationary regressors in VAR models*, Econ. J. 2 (1999), pp. 76–91.
[32] S. Raphael and R. Winter-Ebmer, *Identifying the effect of unemployment on crime*, J. Law Econ. XLIV (2001), pp. 259–283.
[33] G. Saridakis, *Violent crime in the United States of America: A time-series analysis between 1960–2000*, Eur. J. Law Econ. 18(2) (2004), pp. 203–221.
[34] C. Smith and J. Allen, *Violent crime in England and Wales*, Home Office Online Report 18/04, Home Office, London, 2004.
[35] R. Witt and A. Witte, *Crime, prison, and female labor supply*, J. Quant. Criminol. 16(1) (2000), pp. 69–85.
[36] K. Wolpin, *An economic analysis of crime and punishment in England and Wales, 1847–1967*, J. Polit. Econ. 86(5) (1978), pp. 815–840.
[37] D. Wu, *Alternative tests of independence between stochastic regressors and disturbances*, Econometrica 41 (1973), pp. 733–750.

Appendix 1. Cointegration with a mixture of $I(1)$ and $I(0)$ regressors

The error correction representation of the ARDL model can be written as follows:

$$\Delta y_t = c_0 + \pi_1 y_{t-1} + \pi_2 x_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \sum_{k=0}^{q} \delta_k \Delta x_{t-k} + \varepsilon_t, \tag{A1}$$

where y_t is the dependent variable, x_t a vector of regressors, c_0 the drift component and ε_t white noise errors. The approach consists of testing for any long-run relationship between y_t andx_t, that is, the exclusion of the lagged level variables from Equation (A1). The null hypothesis of no cointegration

$$\mathrm{H}_0 : \pi_1 = \pi_2 = 0$$

is tested against the alternative

$$\mathrm{H}_1 : \pi_1 \neq 0, \quad \pi_2 \neq 0$$

using the F-test. Pesaran *et al.* [27] provide two sets of critical values for this statistic covering various specifications of the deterministic terms: The first set assumes that the forcing variables $\{x_t\}$ are $I(0)$ and the second assumes that $\{x_t\}$ are $I(1)$. These two sets provide lower and upper ‘critical value bounds’ covering all possible classifications of $\{x_t\}$ into $I(0)$, $I(1)$ and mutually cointegrated process. If the calculated F-statistic lies above the upper level of the band, the null is rejected. If it falls below the lower level of the band, the null cannot be rejected. If, however, it falls within the band, the result is inconclusive.