Mathematics and Computers in Simulation 33 (1992) 519–532
North-Holland

MODELLING IN ECONOMETRICS: THE DETERRENT EFFECT OF CAPITAL PUNISHMENT

Michael McAleer
Department of Economics
University of Western Australia

1. INTRODUCTION

Econometrics is concerned with the application of statistical or mathematical methods to the analysis of economic data (see, for example, the valuable survey in [1]). Econometric modelling has as its foundation the use of various econometric methods and techniques in evaluating the empirical content of economic relationships. The primary sources of information for use in econometric modelling are economic theory, data, a measurement system, and alternative models for explaining the data. While a notion of "truth" is deemed essential in many scientific studies, it is unhelpful to apply the scientific methods of the natural and physical sciences uncritically to economics. Econometric modelling is not intended to arrive at the truth, but at an adequate representation of non-experimental economic data.

A model specification is the set of all assumptions which defines the parameter space for purposes of inference. It is, therefore, an abstraction from reality. Differences in formulating a model specification may arise through: (a) different theoretical paradigms; (b) different ways of specifying auxiliary assumptions within a paradigm; or (c) different strategies that may be adopted in the process of model construction. Economic theory often indicates which variables enter into a relationship, but it is frequently not so forthcoming regarding the appropriate functional form or other assumptions which define the parameter space. For these reasons, a specification search is warranted, where by a specification search is meant the set of procedures followed in moving from an initial to a final model specification. It is, therefore, a research strategy of examining a number of specifications within a modelling cycle of specification, estimation and evaluation. Conflicting views may, of course, arise regarding the justification of the assumptions and the robustness of the inferences to altering the assumptions.

Testing procedures, especially the use of diagnostics, are the most common specification search method used in econometrics. Special emphasis has been directed towards developing methods for checking the model specification against the data. There has been debate in the literature as to the appropriate methodology to be adopted for econometric modelling (see, for example, [2, 3, 4] and the readings in [5]). Although there are dissenters, a consensus seems to have been reached regarding the essential role of diagnostic tests in the evaluation of econometric models.

As an illustration, consider the linear regression model given by

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + u_t, \quad u_t \sim D(0, \sigma_u^2). \tag{1}$$

Defining the information set $I_t = \{Y_{t-1}, Y_{t-2}, \ldots; X_t, X_{t-1}, X_{t-2}, \ldots; Z_t, Z_{t-1}, Z_{t-2}, \ldots\}$, the key assumptions inherent in estimating equation (1) by ordinary least squares and in evaluating the adequacy of the model are as follows.

(a) Correct functional form: $E(Y_t | I_t) = \beta_0 + \beta_1 X_t + \beta_2 Z_t$ for all $t = 1,2,\ldots,n$.

(b) No heteroscedasticity: $E\{[Y_t - E(Y_t|I_t)]^2 | I_t\} = \sigma_u^2$ for all $t = 1,2,\ldots,n$.

(c) No serial correlation: $E(u_t|I_t) = 0$ for all $t = 1,2,\ldots,n$.

(d) Explanatory variables are exogenous: $E(u_t X_s) = 0$ and $E(u_t Z_r) = 0$ for all t,s and r.

(e) Normality of the errors: $u_t \sim N(0, \sigma_u^2)$ for all $t = 1,2,\ldots,n$.

(f) Structure is constant: β_0, β_1, β_2 and σ_u^2 are constant.

(g) Stationarity: the means, variances and covariances of each of Y_t, X_t and Z_t are constant.

(h) Non-nested models: the model of equation (1) is adequate in the presence of alternative non-nested models.

(i) Robustness/Lack of sensitivity: the model of equation (1) is not unduly affected by alterations in the assumptions.

Although the diagnostic tests *may* suggest alternatives for checking the adequacy of the model, a warning about never losing sight of the alternative is reinforced by taking a glance at Table 1. For any test of the null, there are several possibilities as to the cause of the null hypothesis being rejected, in addition to the specific cause being tested. It is, therefore, essential to keep the options open.

What can be done about reducing the number of possible causes of rejection? Several of these problems are interrelated, and the tests are, in general, not independent. A serious problem arises when any one of the conditions in (a) - (i) (excluding (e)) is not satisfied, since the estimates of the parameters of equation (1) will be inefficient, or the inferences will not be valid, or both (for further details, see [6]). With regard to the quality of the tests employed in

Table 1

Diagnostic Tests for Possible Model Misspecification

Test of Null Hypothesis	Possibilities for Rejecting the Null Hypothesis
(a) **Correct functional form**	(1) incorrect functional form
	(2) omitted variables
	(3) incorrect transformation of variables
	(4) incorrect imposition of restrictions
	(5) serial correlation
	(6) heteroscedasticity
	(7) structural change
	(8) predictive failure
	(9) outliers
(b) **No heteroscedasticity**	(1) additive heteroscedasticity
	(2) multiplicative heteroscedasticity
	(3) omitted variables
	(4) incorrect functional form
	(5) incorrect transformation of variables
	(6) varying coefficients
	(7) incorrect dynamics
	(8) bilinear models
	(9) structural change
	(10) predictive failure
	(11) outliers
(c) **No serial correlation**	(1) autoregressive or moving average errors
	(2) linearly correlated errors
	(3) bilinearly correlated errors
	(4) non-linearly correlated errors
	(5) omitted variables
	(6) incorrect functional form
	(7) incorrect dynamics
	(8) incorrect transformation of variables
	(9) structural change
(d) **Exogeneity**	(1) endogenous explanatory variables
	(2) measurement errors
	(3) serial correlation and lagged dependent variables

Table 1 (continued)

Diagnostic Tests for Possible Model Misspecification

Test of Null Hypothesis	Possibilities for Rejecting the Null Hypothesis
(e) **Normality**	(1) skewness (2) kurtosis (3) other non-normal distributions (4) outliers
(f) **Parameter constancy**	(1) structural change (2) predictive failure (3) outliers (4) heteroscedasticity (5) varying coefficients
(g) **Stationarity**	(1) trending data (2) unit root (3) incorrect transformation of variables
(h) **Non-nested models**	(1) incorrect model (2) incorrect restrictions (3) incorrect dynamics
(i) **Robustness/Lack of sensitivity**	(1) incorrect prior restrictions (2) influential observations (3) fragile assumptions (4) structural change (5) predictive failure (6) outliers (7) choice of sample size (8) choice of sample period (9) choice of instruments (10) inefficient method of estimation

the modelling exercise, is the adequacy of the final specification independent of the procedure by which it was obtained? Unfortunately, it may be taken as being independent only by (foolishly) ignoring the effects of repeated testing and model respecification on the overall level of significance and on the loss of degrees of freedom.

Thus, some problems associated with the use of diagnostic tests lie in an inability to provide clear answers to the following questions.

(1) How many tests should be used?

(2) Are the tests independent?

(3) What ordering of tests is optimal?

(4) Can the tests be used jointly?

(5) How many non-nested alternatives should be considered?

(6) What are the small sample properties of the tests?

(7) Are the tests powerful?

(8) Are the tests robust?

2. AN ILLUSTRATIVE CRIMINOMETRIC EXAMPLE

> **"You can, for example, never foretell what any one man will do, but you can say with precision what an average number will be up to. Individuals vary, but percentages remain constant. So says the statistician."**
> **Sherlock Holmes to Dr. Watson in The Sign of Four. [7]**

Although the commission of murders pervades virtually all modern societies, can anything legal be done to reduce the murder rate? Executions may be of practical concern to some and be of more theoretical concern to others. In the latter category, [8] examined a demand and supply framework in which murder rates were treated as the supply of offenses and deterrent (or punishment) variables as the prices incurred in committing murders. Rational people are thought to fear the threat of execution more than other forms of punishment, such as incarceration. Thus, it has been argued that potential murderers respond to incentives and disincentives in the marketplace.

The example we have chosen for illustrative purposes is due to [8] and has also been considered in [2, 9, 10]. State-by-state cross section data are available to study the deterrent effect of capital punishment in the United States in 1950. The dependent variable, M = the murder rate, is the FBI estimate of the number of murders per 100,000 residents, by state, for 44 states in 1950 (data for the remaining 4 states were unavailable). As given in [8, p. 749], the state murder rates relate to FBI sample surveys of urban areas covering, on average, forty-five percent of the residential population of the 44 states in 1950. Only wilful and felonious homicides are reported as murder data. Justifiable homicides, such as killings in the course of duty by police officers and killings by citizens of felons in the process of committing a felony, are excluded. [8, p. 744] makes the striking observation that the number of deaths inflicted through police intervention is far in excess of the number of executions in the sample under study. The data do not differentiate murders according to criminal liability so that, in self-defence, for example, a murder is recorded in the official statistics but there is no conviction.

The explanatory variables are classified as Economic, Social and Deterrent, as follows.

(1) Economic Variables (opportunity cost of crime)

W = median family income in 1949

X = fraction of families in 1949 with income less than one-half of the median income

U = state unemployment rate

LF = labour force participation rate

(2) Social Variables (taste for crime)

NW = fraction of the state population non-white

AGE = fraction of the state population aged 15-24 years

URB = fraction urban

SOUTH =dummy variable (=1 for Southern states)

(3) Deterrent Variables (price of committing crime)

PC = probability of conviction (convictions for murder in 1950 divided by the estimated number of murders in the state)

PX = probability of execution conditional on conviction (average number of executions for 1946-50 divided by the number of convictions)

T = length of sentence (median time in months served by prisoners convicted of murder and released in 1951).

In the four papers mentioned above, there was also an executing state dummy variable, namely

XPOS = dummy variable (=1 if PX is positive).

Of the 44 states, 35 had at least one execution between 1946 and 1950, although several states which did not have executions still retained the death penalty. It is worth noting that, in previous research, XPOS has been used as a Deterrent variable and its presence has been found to affect the empirical results significantly. A clearer interpretation of XPOS may be gleaned from a disaggregation of the 48 states according to a common geographical distribution, which is given in [9, p. 102]. Four of the 48 states are excluded from the study through lack of data. The 9 non-executing states generally have low murder rates, whereas many of the executing states have high murder rates. Thus, rather than interpreting the existence of the death penalty as a deterrent, it is argued in this paper that the dummy variable simply captures differential murder rates between executing and non-executing states. It is perhaps also worth noting that all Southern states have executions.

Some difficulties should be noted at the outset. Omitted explanatory variables which could influence M include the degree of ownership of firearms, the efficacy of gun laws (where they exist), the efficiency of law enforcement agencies (which may be poorly approximated by PC), the quality of and accessibility to medical care (which may not be fully captured by URB), and qualitative political variables such as whether a state has a conservative administration and/or constituency. Political factors can play an important role since PC, PX and T are subject to the control of public authorities, that is, these variables may be sensitive to changes in M instead of the reverse. Indeed, PX may itself be explained by M in past periods, as well as overcrowding in prisons, the cost of incarceration, and the conservatism of the judiciary. In our empirical example, the data for M relate to 1950 while data for PX relate to the period 1946-50. Thus, M is unlikely to have influenced PX, but PX may well have influenced M.

The simple linear regression model to be estimated specifies M as the dependent variable which is to be explained by some or all of the explanatory variables (or regressors). All the empirical results are based on the standard heteroscedasticity - grouping transformation of multiplying the dependent and all explanatory variables by the square root of the FBI survey population. Estimates of the general specification, which includes all the explanatory variables available, are given in the first column of Table 2. Apart from the sign of U, which appears to be incorrect but insignificant, all estimated coefficients appear to have the correct signs. The regressor LF is ambiguous, since an increase in LF may have a positive effect if it represents a breakdown in the traditional (law-abiding) family unit, or negative if it leads to greater accessibility to employment opportunities. However, owing to collinearity among the many regressors, only XPOS (which has a positive effect) and PC and PX (which have negative effects) are statistically significant. The $\bar{R}^2$ value and the F-statistic indicate the regression is useful,

Table 2

Estimates of the General Specification and Three Acceptable Prior Beliefs[1]

Regressors	General	Rational Maximizer	Crime of Passion	Perverse
Intercept	-13.519 (0.907)	-6.597 (0.586)	-24.930 (1.577)	1.743 (0.243)
XPOS	4.263** (3.280)	6.246** (5.224)	2.998* (2.337)	3.474** (2.874)
W	-0.003 (1.523)	-0.007** (5.862)	-0.003 (1.328)	
X	9.959 (0.595)	-1.971 (0.153)	34.108* (2.066)	
U	-5.851 (0.235)	-28.190 (1.195)	-28.205 (1.278)	
LF	32.411 (1.565)	53.953* (2.646)	28.150 (1.219)	
NW	8.400 (1.838)		9.887 (1.979)	10.388* (2.367)
AGE	27.540 (0.567)		30.700 (0.576)	34.924 (0.797)
URB	0.004 (0.088)		0.053 (1.228)	-0.021 (0.789)
SOUTH	2.791 (1.839)		4.000** (2.813)	3.674** (2.873)
PC	-7.705* (2.242)	-7.231* (2.070)		-8.782** (2.908)
PX	-20.097* (2.467)	-20.509* (2.446)		-22.435** (2.990)
T	-0.014 (1.836)	-0.025** (3.622)		-0.015* (2.220)

* denotes significant at 5%.
** denotes significant at 1%.

Table 2 (continued)

Estimates of the General Specification and Three Acceptable Prior Beliefs[1]

Descriptive Statistics	General	Rational Maximizer	Crime of Passion	Perverse
$\bar{R}^2$	0.842	0.814	0.802	0.840
SE	6.528	7.077	7.305	6.573
F-statistic	20.101**	24.579**	20.377**	29.190**

Diagnostic Tests	General	Rational Maximizer	Crime of Passion	Perverse
Functional Form	4.121	2.962	2.910	5.453*
Normality	0.845	0.289	1.039	0.889
Heteroscedasticity	1.617	0.441	2.382	2.677

[1] The dependent variable throughout is M. Absolute t-ratios are given in parentheses. $\bar{R}^2$ is the adjusted squared multiple correlation coefficient, SE is the standard error of the regression and the F-statistic tests the significance of the regression equation. The diagnostic tests are for functional form misspecification (Ramsey's RESET test), normality (asymptotic chi-squared test), and heteroscedasticity (asymptotic chi-squared test).

* denotes significant at 5%.
** denotes significant at 1%.

while the three diagnostic tests suggest the estimated general model is statistically adequate regarding functional form, normality and homoscedasticity.

Alternative nested and non-nested models may arise through alternative prior beliefs (or broad stereotypes). Five priors available in the literature, together with three new prior beliefs, are presented below. The Right Winger believes that murder can be explained only by punishment, and so includes the Deterrent variables but excludes the Economic and Social variables. An Eye-for-an-eye prior holds that murder can be explained only by capital punishment, which must be preceded by lawful conviction. Thus, PC and PX are the included Deterrent variables, whereas the Deterrent variable T is excluded. Unlike the previous prior, Rambo does not need the luxury of conviction prior to execution, and argues (violently) that only PX is required. The Rational Maximizer balances the potential economic gains against the potential costs of punishment, and is therefore affected by both the Deterrent and Economic variables. The Bleeding Heart prior views murder as the product of economic deprivation and impoverishment, so that only Economic variables are included. A Crime of Passion prior believes that a person who commits murder is not considering possible punishment, but is a product of Economic and Social conditioning. An Abolitionist argues that capital punishment is barbaric in a civilized society, and omits only PX while including all other variables. Finally, a Perverse (or "Non-Economic") prior does not view Economic variables as leading to murder, but considers the Deterrent factors and Social influences as being important.

Nested tests are performed for each of the seven prior beliefs (that is, tests of the validity of the excluded variables) against the general specification, and the results are presented in Table 3. The Right Winger, Eye-for-an-eye, Rambo and Bleeding Heart priors are rejected convincingly, implying that each of these prior beliefs excludes statistically significant explanatory variables. In the four remaining cases, all the priors are acceptable at the 1% level of significance, so the decisions to exclude different sets of variables are acceptable to the data.

The outcomes of further representative nested tests of one prior against a more general prior are presented in Table 4. For example, the Rational Maximizer includes three Deterrent variables in addition to the four Economic variables of the Bleeding Heart, in which case the Bleeding Heart is the null model to be tested against the more general alternative Rational Maximizer model. The outcome of the test is to reject the null at the 1% level of significance, so that the Rational Maximizer prior is preferred. In other cases, the Rational Maximizer is also preferred to the Right Winger, which is also dominated by the Perverse prior, while the Crime

Table 3

Nested Tests of Prior Beliefs Against the General Specification[1]

Prior Beliefs	Variables Included	Variables Excluded	F-statistic	Decision to Exclude Variables
Right-Winger	Deterrent	Economic and Social	F(8,31)=9.106	Unacceptable
Eye-for-an-eye	PC and PX	T, Economic and Social	F(9,31)=17.140	Unacceptable
Rambo	PX	PC, T, Economic and Social	F(10,31)=16.178	Unacceptable
Rational Maximizer	Deterrent and Economic	Social	F(4,31)=2.533	Acceptable at 5%
Bleeding Heart	Economic	Deterrent and Social	F(7,31)=4.808	Unacceptable
Crime of Passion	Economic and Social	Deterrent	F(3,31)=3.858	Acceptable at 1%
Abolitionist	PC, T, Economic and Social	PX	F(1,31)=6.086	Acceptable at 1%
Perverse	Deterrent and Social	Economic	F(4,31)=1.120	Acceptable at 10%

[1] The full set of variables includes Deterrent, Economic and Social. An intercept term and the executing state dummy variable are included in each model.

of Passion prior dominates both the Abolitionist and Bleeding Heart priors. Thus, the three prior beliefs not rejected by the data are the Rational Maximizer, Crime of Passion and Perverse, the results of which are presented in the last three columns of Table 2. All but two of the signs of the Rational Maximizer model are as expected, with the remaining two (X and U) being insignificant. Only U yields an incorrect sign for the Crime of Passion model, but it is insignificant. All signs for the Perverse model are as expected. The executing state dummy variable is significant in all three models, and all the Deterrent variables are significant in both the Rational Maximizer and Perverse models. On the basis of $\bar{R}^2$ values, the Perverse model fits the data best, and almost as well as the general model. At the 1% level of significance, all three diagnostic tests indicate the models are statistically adequate.

Table 4

Tests of Nested Specifications Against Acceptable Priors[1]

Null Model	Alternative Model	F-statistic	Decision
Bleeding Heart	Rational Maximizer	$F(3,35)=6.673$	Reject null at 1%
Right Winger	Rational Maximizer	$F(4,35)=13.341$	Reject null at 1%
Right Winger	Perverse	$F(4,35)=16.859$	Reject null at 1%
Crime of Passion	Abolitionist	$F(2,32)=2.368$	Do not reject null
Bleeding Heart	Crime of Passion	$F(4,34)=4.409$	Reject null at 1%

1 An intercept term and the executing state dummy variable are included in each model.

Non-nested test statistics are calculated for testing the following paired models: Rational Maximizer versus Crime of Passion, Rational Maximizer versus Perverse, and Crime of Passion versus Perverse. Six non-nested procedures are used in the Microfit econometric package (see [11] for further details), and the results are presented in Table 5. Taking the first named model as the null in each case, the decisions are as follows: the Rational Maximizer is rejected by Crime of Passion, and vice-versa; the Rational Maximizer is rejected by Perverse but Perverse is not rejected by the Rational Maximizer; and Crime of Passion is rejected by Perverse but Perverse is not rejected by Crime of Passion. Thus, the only model regarded as statistically

Table 5

Tests of Non-nested Specifications[1]

Test Statistic	Ratinal Maximizer against Crime of Passion	Crime of Passion against Rational Maximizer
N	-4.779**	-5.339**
NT	-2.513*	-3.153**
W	-2.225*	-2.724**
J	3.112**	3.278**
JA	1.870	1.843
F	2.533	3.858*
(Decision)	(Reject null)	(Reject null)

Test Statistic	Ratinal Maximizer against Perverse	Perverse against Rational Maximizer
N	-4.707**	-2.419*
NT	-3.266**	-1.566
W	-2.647**	-1.407
J	3.285**	1.966*
JA	2.934**	1.467
F	2.533	1.120
(Decision)	(Reject null)	(Do not reject null)

Test Statistic	Crime of Passion against Perverse	Perverse against Crime of Passion
N	-8.587**	-2.357*
NT	-4.127**	-0.861
W	-3.517**	-0.825
J	3.508**	1.901
JA	3.365**	0.769
F	3.858*	1.120
(Decision)	(Reject null)	(Do not reject null)

1 An intercept term and the executing state dummy variable are included in each model.
* denotes significant at 5% and rejection of the null (that is, first named) model.
** denotes significant at 1% and rejection of the null (that is, first named) model.

adequate after the non-nested tests is based on the Perverse prior belief. As seen from Table 2, the deterrent effect of capital punishment appears to be supported strongly by the data for the dominant Perverse prior.

ACKNOWLEDGEMENTS

This paper is based on part of my inaugural lecture at the University of Western Australia, and was presented as a keynote address to the 9th Biennial Conference on Modelling and Simulation, Gold Coast, Australia, December 1991. It developed from a graduate course on Econometric Methodology I taught for several years in the Department of Statistics at the Australian National University. I wish to thank Tony Jakeman, Colin McKenzie, Les Oxley, Peter Phillips and seminar participants at Bond University, Griffith University and the University of Edinburgh for helpful comments, and to acknowledge the financial support of the Australian Research Council.

References

[1] Pesaran, M.H. (1990), Econometrics, in Eatwell, J., Milgate, M. and Newman, P. (eds.), The New Palgrave: Econometrics (Macmillan, London), 1-34.
[2] Leamer, E.E. (1983), Let's take the con out of econometrics, American Economic Review, 73, 31-43 (reprinted in [5, 29-49]).
[3] McAleer, M., Pagan, A.R. and Volker, P.A. (1985), What will take the con out of econometrics?, American Economic Review, 75, 293-307 (reprinted in [5, 50-71]).
[4] Pagan, A.R. (1987), Three econometric methodologies: a critical appraisal, Journal of Economic Surveys, 1, 3-24 (reprinted in [5, 97-120]).
[5] Granger, C.W.J. (ed.) (1990), Modelling Economic Series: Readings in Econometric Methodology, Oxford University Press, Oxford.
[6] Fiebig, D.G., McAleer, M. and Bartels, R. (1991), Properties of ordinary least squares estimators in regression models with non-spherical disturbances, to appear in Journal of Econometrics.
[7] Conan Doyle, A. (1984), The Penguin Complete Sherlock Holmes (Penguin, Harmondsworth).
[8] Ehrlich, I. (1977), Capital punishment and deterrence: some further thoughts and additional evidence, Journal of Political Economy, 85, 741-788.
[9] McAleer, M. and Veall, M.R. (1989), How fragile are fragile inferences? A re-evaluation of the deterrent effect of capital punishment, Review of Economics and Statistics, 71, 99-106.
[10] McManus, W.S. (1985), Estimates of the deterrent effect of capital punishment: the importance of the researcher's prior beliefs, Journal of Political Economy, 93, 417-425.
[11] Pesaran, M.H. and Pesaran, B. (1989), Microfit: An Interactive Econometric Software Package (Oxford University Press, Oxford).