Forecasting Residential Burglary*

Derek Deadman

Public Sector Economics Research Centre
Department of Economics
University of Leicester
February 16, 2000

Tel: 0116 2522885
Fax: 0116 2522908

Abstract

Following the recent work of Dhiri et al (1999) at the Home Office predicting recorded burglary and theft for England and Wales to the year 2001, econometric and time series models have been constructed for predicting recorded residential burglary to the same date. A comparison between the Home Office econometric predictions and the less alarming econometric predictions made in this paper identifies the differences as stemming from the particular set of variables used in the models. However, these econometric models adopt an error-correction form which appears in both cases to be the main reason why the models predict increases in burglary. To identify the role of error-correction in these models, time series models have been built for the purpose of comparison, all of which predict substantially lower numbers of residential burglaries. The next three years would appear to offer an opportunity to test the utility of error-correction models in the analysis of criminal behaviour.

Keywords: Residential Burglary: error-correction; time series forecasting

* I am grateful to participants at the Home Office 'Trends in Crime' seminar, held in December 1999, for their helpful and constructive comments.

1. Introduction

The number of recorded residential burglaries in England and Wales has been declining since 1993. Such a decline is neither unique to this category of crime nor just limited to England and Wales, as similar patterns in recorded crime over this period have been experienced in many other European countries and in the United States (see Field (1999), p.16). Whilst there are now several published econometric analyses of recorded crime, until recently only one (Deadman and Pyle (1997)) appears to have been used for prediction. It is therefore most encouraging that a significant break with past practice has been made with the publication of Home Office Research Study No.198 (Dhiri et al (1999)) specifically addressing the issue of modelling and predicting property crime trends.

The Home Office predictions have been made on the basis of forecasts of demographic and economic changes and on the assumption that no other factors (such as government policies to reduce crime) are altered over the prediction period. It is predicted that the recent declines in recorded property crime will be reversed from 1999 onwards and the level of such crime will rise substantially (increases of 25% for burglary and 40% for theft compared with 1997) by 2001. These are dramatic and brave predictions, which are clearly of deep concern to politicians making policy in this area and to Chief Constables who may be required to meet targets set for them by policy makers. Such predictions are also of considerable interest to the general public and they did receive widespread press coverage on release.

It will be widely appreciated that the basis for the predictions produced at the Home Office has now been made public and is available for comment and analysis. Together with the predictions presented below, the Home Office predictions will provide a powerful test of the value of the particular type of econometric forecasting model (termed an 'error-correction' model) which has come to be increasingly used for empirical research in this area. Accordingly, for comparative purposes, also presented are sets of forecasts based on traditional time-series methodology.

The econometric model used to make predictions is based on that reported in Pudney et al (2000) for residential burglary but with the addition of police numbers within the set of explanatory variables. Further discussion of this model may be found in Macdonald and Pyle (2000, Chapter 2). Residential burglary rather than burglary as a whole has been investigated, as it is for this category that independent information exits for the degree of under-recording in the British Crime Survey and the General Household Survey. It was the potential bias due to under-recording that was the main focus of the work of

Pudney et al (2000). Residential burglary represents just over half of all recorded burglary offences. Previous work (e.g. Pyle and Deadman (1994), Deadman and Pyle (1997), Hale (1998)), including that of the Home Office (Field (1999), Dhiri et al (1999)), suggests that all recorded burglary could be modelled on very similar lines to residential burglary with results that would be comparable to those found for residential burglary. The time series patterns for recorded residential burglary and all recorded burglary are very similar. For the sample period 1946 to 1997, the correlation between the two series was 0.997.

2. Econometric Analysis

The estimated model used here for prediction of the number of recorded residential burglary offences includes economic activity variables (consumption and unemployment), criminal justice variables (probability of conviction, probability of imprisonment, length of sentence and number of police) and a demographic variable (number of males aged 15-24). The model was estimated in natural logarithmic form using annual data from 1950 to 1997 and incorporated a dummy variable to take account of a change in recording practice following the Theft Act of 1968.

In contrast, the Home Office model predicts separately both theft and burglary using just two explanatory variables, namely the stock of goods (proxied in each year by the sum of total household final consumption expenditure in the current and three preceding years) and the number of males aged 15 and 20. The theft and burglary series were adjusted to take account of the new counting rules introduced in the Theft Act. Annual data for 1951 to 1998 were used in the estimation of the model which was also in logarithmic form.

What is common to the econometric predictions in this paper and those presented by the Home Office is that they are both obtained from models incorporating an error-correction term. In both cases crime is modelled as having a long run equilibrium solution together with a mechanism which allows for dynamic adjustment to this long run path from positions off this path. This structure appears to be of central importance in the pattern of predictions discussed below, which were obtained from the following model:

TABLE 1

Ordinary Least Squares Estimation

Dependent Variable is Δ Resburg

All variables in natural logarithms

	Coefficient	t-ratio	P-value
Δ unemployment	.26302	4.2472	0.000
Δ consumption	-1.7459	-3.5393	0.001
Δ conviction	-0.3452	-2.5156	0.018
Δ sentence	-0.3536	-1.9378	0.062
Δ imprisonment	-0.2484	-2.0517	0.049
Δ police	-1.4946	-1.7465	0.091
Δ youths	0.7336	0.8024	0.429
Δ dummy	0.4671	6.3789	0.000
Resburg(-1)	-0.2426	-3.1569	0.004
Unemployment(-1)	0.1060	1.9761	0.058
Consumption(-1)	0.9458	3.0968	0.004
Conviction(-1)	-0.1317	-0.9314	0.359
Sentence(-1)	-0.6144	-2.9047	0.007
Imprisonment(-1)	-0.0757	-0.7028	0.488
Police(-1)	-1.7321	-2.9626	0.006
Youths(-1)	0.8060	3.2374	0.003
Dummy	0.1487	2.3062	0.028
Intercept	-2.0640	-0.8205	0.419

Notes:

47 Observations used for estimation from 1951 to 1997.
All variables in natural logarithms.

R^2 = 0.92834	R-Bar-Squared = 0.88633
S.E. of Regression = 0.47823	F-Stat. F(17,29) 29.0992 (.000)
Mean of Dep Var = 0.059911	S.D. of Dep Var = 0.14185
RSS = 0.066324	Equation Log-likelihood = 87.5488
Akaike Info. Criterion = 69.5488	Schwarz Bayesian Criterion = 52.8974
DW Statistic 1.9499	

Serial Correlation	c^2(1) = 0.023951 (.877)	F(1,28) = 0.014276 (.906)
Functional Form	c^2(1) = 0.064493 (.800)	F(1,28) = 0.038474 (.846)
Normality	c^2(2) = 2.7805 (.249)	Not Applicable
Heteroscedasticity	c^2(2) = 0.46509 (.495)	F(1,45) = 0.44975 (.506)

Definitions of the variables used are given in the Appendix. The model was estimated by Ordinary Least Squares following the approach of Sims, Stock and Watson as discussed in Pudney et al (1997). This model provides a good fit to the sample data, and passes all the standard diagnostic tests. The estimated coefficients of the criminal justice variables indicate a significant deterrence role for these variables, and both consumption and unemployment appear to have some power in the explanation of residential burglary.

For prediction, some assumptions need to be made regarding the values of the explanatory variables outside of the sample period. These assumptions were as follows:

Assumption 1. All criminal justice variables (conviction rate, probability of imprisonment, sentence length, number of police) were set at their values in 1997.

Assumption2. Population projections (both for totals and for the number of males aged 15-24) were taken from GAD (1999).

Totals:

UK	1996 (base)	58,801,000	2001	59,618,000
England and Wales	1996 (base)	52,010,000	2001	52,818,000

Male Youths:

England and Wales	1996 (base)	3,290,000	2001	3,297,000

Assumption 3. Forecasts for consumption expenditure are Treasury forecasts taken from HMT (1999).

Household consumption (percentage change from previous year)

1998	2.25%	1999	1 to 1.5%
2000	2.75%	2001	2.75 to 3.25%

Assumption 4. Forecasts for unemployment are those made by the National Institute of Economic and Social Research (NIESR, (1999)).

Unemployment (Claimant Count)

1998	1,347,000	1999	1,253,000
2000	1,234,000	2001	1,282,000

Whilst the demographic forecasts suggest only a small rise in the total number of males aged 15-24 years over the forecast period, this conceals a predicted 6% rise in the 15-19 year old age band compared to a 5% fall in the 20-24 year band. If the former age group has a higher propensity to commit residential burglary, the forecasts from this model will tend to understate the demographic effect.

The dynamic forecasts for residential burglary which result from the use of the assumptions and values above for the explanatory variables are given in Table 2 with the Home Office projections for aggregate Burglary in Table 3. The forecasts reported here and in later tables are made under the 'old rules' operated by the police for counting of offences rather than the new rules introduced in 1998.

TABLE 2

Residential Burglary: Projections 1998-2001

(Plus/Minus Two Standard Errors of Forecast)

	Lower	Median	Upper	Annual Change (%)
1996		602128		
1997		519265		-13.8
1998	437689	496416	563018	-4.4
1999	414728	504326	613281	+1.6
2000	399703	516731	668023	+2.5
2001	401317	552002	759266	+6.8

TABLE 3
Home Office: HORS 198
Burglary: Projections 1998-2001

	Number (millions)	Annual Change (%)
1997	1.02 (actual)	
1998	0.97 (actual)	-5.0
1999	1.02	+6.0
2000	1.14	+11.0
2001	1.28	+12.0

The Home Office also used its model to predict theft to 2001. A similar pattern of predictions to those for burglary was reported, but with theft predicted to rise at an even faster rate than burglary (40% more recorded offences by 2001 compared with 1997). Dhiri et al (1999, p.20) do present some reasons why their predictions may overstate what will happen under the assumptions they have taken, but very substantial rises of at least 17% in both recorded theft and burglary between 1997 and 2001 are still predicted. This is virtually three times the increase suggested by the model presented here and it is interesting to consider why such a difference exists between the two sets of predictions.

Both models incorporate current household consumption expenditure and a demographic variable. Inspection of the forecasts used for these explanatory variables alone would not lead one to expect the observed tendency of both models to predict a reversal of recorded property crimes from 1999 and an increasing annual rate of growth in such crime to 2001. In both models, it appears that this pattern is driven by the presence of an error-correction term. The rapid growth (above trend) in household consumption expenditure towards the end of the 1990s has pushed recorded crime below its long run equilibrium values and the forecast values represent the model returning to these values. Indeed, it was this feature of recorded crime series noted by Field (1990) that prompted the use of error-correction models by Pyle and Deadman (1994). However, there are very great quantitative differences in the two sets of forecasts. Model specifications are markedly dissimilar, even though superficially they incorporate similar variables and are estimated over virtually identical sample periods. In our specification, even though criminal justice variables are taken as unchanging for the forecast periods, their presence in the estimated equation will affect parameter estimates, as will the presence of the unemployment variable. Neither unemployment nor criminal justice variables appear in the Home Office models. Instead, much of the explanation of past and future crime trends appears to reside in the cumulated consumption variable, taken to represent the stock of consumer goods available for theft or burglary. Given the large stock of consumer durables in existence at any one time and the relatively small part of it that is stolen, it is difficult to believe that, over the sample period, any burglar would have had his or her criminal intentions frustrated by a lack of desirable goods to steal. The behavioural underpinnings of this stock effect appear weak.

3. Time Series Analysis

One would not expect traditional time-series forecasting models, such as Box-Jenkins models, to replicate the above predicted patterns for recorded crime given the assumptions made about the future state of the economy (consumption and unemployment) and the relatively small forecast increase in the number of males aged 15 to 24 over the forecast period. Stationary univariate Box-Jenkins (ARIMA) models produce optimal (minimum mean squared error) forecasts that revert quickly to the mean of the process, which are therefore only intended for short run forecasts. However, such models are useful for obtaining an initial specification of the noise component of multivariate transfer function models which allow for the influence of independent variables and hence are available for longer run forecasting.

The autocorrelation function (acf) of the log of residential burglary shows a clearly nonstationary series and the time series plot reveals the structural (recording) break in 1968. Subsequent modelling is conducted in first differences of the variable (which are stationary) and utilising a dummy variable to account for the break, replicating in both respects actions taken in the econometric analysis discussed above. The standard rounds of identification, estimation and diagnostic checking were conducted on a series of univariate models for the first difference of the natural logarithm of residential burglary, yielding the following parsimonious first order autoregressive model as an adequate description of the stochastic process underlying residential burglary:

$$\Delta \text{Resburg}_t = \underset{(4.07)}{0.535}\Delta \text{Resburg}_{t-1} + \underset{(8.78)}{0.722}\Delta \text{dum}_t + \text{residual}.$$

where student-t values of the coefficients are given below in parentheses. The residual mean square for this model was 0.00867. The acf and pacf (partial autocorrelation function) of the residuals of this estimated equation revealed no significant residual autocorrelations. Additionally, the joint test for the significance of the first m residual autocorrelations given by the Ljung-Box Q statistic (asymptotically distributed as c^2 with m-p degrees of freedom where p is the number of estimated model parameters) produced a statistically insignificant value of 13 at lag 20 at the 5% level of significance. The addition of further autoregressive terms (overspecification test) gave no statistical improvement over the model given above, with insignificant coefficients on higher lagged variables and larger residual mean squares.

For comparison purposes, this univariate model gave the following forecasts:

TABLE 4
Residential Burglary: Univariate Projections 1998-2001 from a AR(1) process
(Plus/Minus Two Standard Errors of Forecast)

	Lower	Median	Upper	Annual Change (%)
1998	398717	480058	578716	-7.6
1999	328085	461504	649181	-3.9
2000	279078	451516	730502	-2.2
2001	244638	448228	821245	-0.7

Forecast error bands which widen as the forecast horizon lengthens are a feature of optimal forecasts from nonstationary models. In this model, the intervals are larger than those given by the econometric model discussed earlier. The forecast changes converge to a zero mean as expected from a model with no significant constant term.

Given existing knowledge on practical identification procedures for transfer function modelling (especially where the input variables are expected to be intercorrelated), it seems advisable to estimate transfer function models employing only a restricted range of the variables employed in the econometric estimation. In order to make the results as comparable as possible to those produced by the Home Office, consumption, unemployment and the youths variable have been used. Following the methodology first proposed by Box and Jenkins (1970) (see also McLeod (1982) and Vandaele (1983)), separate univariate models were built for each of the independent variables. These were used for prewhitening the output variable in the identification stage of modelling in a 'piece-meal' fashion to specify the complete transfer function model. This approach may be expected to work quite well in the multiple input case provided the independent variables are only weakly related (see Mills (1990, p.261). As each variable is differenced for stationarity, this requirement is met, with the sample correlations of the differences of the natural logarithms being (1951 to 1997):

Δconsumption and Δyouths 0.0666

Δconsumption and Δunemployment -0.365

Δyouths and Δunemployment 0.299

Transfer function modelling is particularly suited to situations such as that considered here where there is one way causation between inputs and the output variable, with no possibility of feedback effects. Alterations in consumption, unemployment and the age structure variable may have contemporaneous or future effects on residential burglary, but will not themselves be affected by residential burglary. Hence, when using cross correlation functions between prewhitened input and output variables for model identification, only the pattern of cross correlation coefficients at zero and positive lags will be of interest.

The univariate model for the first differences of the logarithm of consumption which was selected using the standard diagnostic tests outlined above was a mixed model involving a first order autoregressive term and a second order moving average term. The residuals from this model were cross-correlated with the prewhitened output series obtained by using the model for consumption as the filter. The only statistically significant cross-correlation coefficient was at lag zero, indicating an initial specification of the transfer function model in which the consumption variable has only a contemporaneous effect on residential burglary. The parsimonious choice for the univariate model for the first differences of the logarithm of unemployment was a second order autoregressive model. The cross-correlation function between the residuals of this model and the prewhitened output variable obtained from the use of the unemployment filter had statistically significant coefficients at lags 0 and 2, and a near significant coefficient at lag 3. However, there did not appear to be any evidence of a pattern in the higher order coefficients which would have suggested adopting a rational lag structure for the specification of the transfer function between unemployment and residential burglary. The form of the ccf suggests that the influence of the unemployment variable on crime may be distributed over a longer period than that of consumption. A mixed first order autoregressive moving average model was selected as the parsimonious description for the first differences of the youths variable. Respecifying this model with the addition of further autoregressive terms clearly pointed to parameter redundancy and omitting the moving average term gave rise to significant acf coefficients for the model residuals. However, the ccf of the residuals from the selected model with the prewhitened output had no significant coefficients. This suggests that the youths variable could be omitted from the complete specification of the multiple input transfer function model. This conclusion was confirmed by adding this variable to models including both the consumption and unemployment variables. In these extended models, the youths variable was invariably statistically insignificant, and the resulting estimated model diagnostics were not improved by the presence of this variable.

On the basis of the prewhitening exercises described, an initial transfer model was estimated involving (in differences) the contemporaneous level of consumption, unemployment at lags 0, 1 and 2, and the dummy variable. Two of the important diagnostic checks on model adequacy are those to check that the residuals of the estimated model behave as white noise and that they are uncorrelated with the prewhitened inputs (Vandaele (1983, pp.306-313). The latter requirement was not met by the initial model when the errors of the equation are specified as white noise, with both cross correlation functions displaying significant coefficients at lag zero. The estimated initial model also had statistically insignificant coefficients for the longer lags of the unemployment variable. The misspecification of this part of the transfer function model could be the cause of the non zero cross correlation coefficients, so the model was re-estimated with these variables omitted. The resulting model had a marginally lower root mean square error, but retained the earlier cross correlation problem, and had a near statistically significant autocorrelation function coefficient at lag 1 for the model residuals, indicating a possible misspecification in the noise component of the model. Interestingly for what follows, both the initial transfer function model and this re-specification produced forecasts which indicated an upturn in residential burglary by the end of the forecast period (year 2001).

The statistical weaknesses of the models above were solved by formulating the noise component on the lines of the univariate model for residential burglary discussed earlier. That is, first and second order autoregressive terms for the differences of residential burglary were included in the estimated model, along with differences of contemporaneous consumption and unemployment. This model passes the diagnostic tests applied to it, including a set of residuals which can be accepted as white noise (an LBQ statistic of 12 at lag 20 and no pattern or significant coefficients in their acf or pacf) and which display no statistically significant cross correlation coefficients with prewhitened consumption or unemployment. For purposes of comparison with the error-correction model reported previously, this model was estimated as follows:

TABLE 5
Transfer Function Estimation
Dependent Variable is Δ Resburg
All variables in natural logarithms

	Coefficient	t-ratio
Δ unemployment	0.0909	1.98
Δ consumption	-2.590	-5.06
Δ Resburg (-1)	0.4821	3.20
Δ Resburg (-2)	0.3938	2.61
Δ dummy	0.7017	10.33

RMS = 0.0054; LBQ = 12 at lag 20.

Compared with the univariate model for residential burglary (which had a single autoregressive term), this model has a substantially lower RMS statistic, indicating that the economic activity variables contain information useful to the explanation of residential burglary over and above that contained in its own past history. Addition of further lagged residential burglary terms yields no model improvements, and the exclusion of the highest lagged residential burglary variable leads to a markedly inferior model. The addition of the contemporaneous youths variable to the transfer model above leads to an estimated model which is virtually unaffected in terms of coefficient values, fit or diagnostic tests, and the variable itself has an insignificant coefficient. In this formulation, therefore, there seems to be no reason to include the youths variable. The forecasts which arise from the transfer function model, using the same values for consumption and unemployment as were used in the error-correction model were as follows:

TABLE 6

Residential Burglary: Transfer Function Projections 1998-2001 (Plus/Minus Two Standard Errors of Forecast)

	Lower	Median	Upper	Annual Change (%)
1998	403754	467700	541772	-9.9
1999	332502	432459	562464	-7.5
2000	264084	396489	595278	-8.3
2001	209208	365479	638478	-7.8

The width of the forecast interval increases rapidly as the lead time increases which, as was remarked on earlier, is a feature of these models. This problem was also identified in the Home Office study of Dhiri et al (1999, p.18), and as stated there, one would have to treat forecasts for more than three years ahead as unreliable.

The identification stage of the transfer function modelling exercise involves the use of cross correlation functions between prewhitened independent variables (consumption, unemployment and the youths variable) and prewhitened output (residential burglary) obtained by the use of filters (univariate models) estimated for the independent variables. However, the Theft Act of 1968 resulted in a redefinition of offences such that the number of recorded residential burglaries more than doubled between 1968 and 1969. Although in the final estimated transfer function models this break is adequately captured by the inclusion of a dummy variable, the calculation of the cross correlation functions above did not allow for this break. Accordingly there is a possibility that the sequence of models investigated and hence the final model selected might have been affected by a pattern of potentially 'contaminated' cross correlation coefficients. As the ccfs between prewhitened inputs and prewhitened output were conducted using first differences of the variables, the break would only appear in the output variable as a pulse or outlier at 1969, but this could affect all estimated ccf coefficients.

To investigate this issue, an adjusted residential burglary series was constructed in which an attempt was made to remove the effect of the 1968 Theft Act from the data series from 1969 onwards. A linear trend was fitted to the logarithms of recorded residential burglary per capita using the data up to and including 1968 and the resulting trend line was then used to forecast the value for 1969. The difference between the actual and forecast values for 1969 was taken to represent the effect of the Act and was subtracted from all the data post 1968. This was the same procedure as that adopted when testing for the order of integration of the residential burglary series described in Pudney et al (2000). The adjusted series was then used in the identification stage of modelling to choose the initial transfer function model. A comparison of the ccfs obtained from the use of the original and adjusted series revealed that the adjusted series gave rise to patterns of ccf coefficients which were simpler to interpret. For the ccfs using the adjusted series of the number of recorded residential burglaries, the only statistically significant coefficients in the ccfs were those at zero lags for both the consumption and unemployment variables. This suggests that the use of the adjusted series would have lead to the same chosen final transfer function model more rapidly than the modelling exercise described earlier which used the unadjusted data.

4. Comparison of Prediction Profiles

It is informative to consider what are the essential features of each model which give rise to the observed differences in the forecast profiles. The univariate model is driven by the sign of the last sample difference (negative) in residential burglaries together with a non significant sample mean for these differences. For the transfer function, it appears that the presence of two autoregressive terms, the last sample values of which are both negative, leads to the pattern of declining forecast values. As the forecast values of both the consumption and unemployment values are broadly in line with their values towards the end of the sample period, there is no countervailing force exerted by the presence of these variables in the forecast period. This conjecture appears to be confirmed by the forecasts from transfer function models which exclude the autoregressive terms where it is only at the end of the forecast period (where unemployment is predicted to rise) that residential crime is forecast to rise. However, as noted above, transfer function models which exclude autoregressive terms have to be judged as inferior representations of the process for residential burglary on the grounds of diagnostic checks.

The one type of model which predicts a reversal of the recent downward trend in residential burglaries is the error-correction model. It is this structure which underlies the Home Office model (Dhiri et al, 1999) and it is the return of the number of recorded burglaries to 'underlying levels' which appears to be the basis of their predicted increases for the years to 2001. This is also true of the error correction model estimated above. The presence of criminal justice variables affects the values of estimated parameters on the consumption and unemployment variables and may, in part, account for the lower forecast increases from this model compared with those from the Home Office. However, if the error-correction terms are excluded from the model specification and a purely 'short-run' model in differences is estimated, then using the same forecast values for the criminal justice, consumption and unemployment variables as before, the following estimated model and associated forecasts are obtained:

$$\begin{aligned}\Delta Resburg_t = \; & \underset{(2.11)}{0.134\Delta Unem_t} - \underset{(-0.624)}{0.256\Delta Consum_t} - \underset{(-5.34)}{0.710\Delta Convict_t} - \underset{(-0.976)}{0.249\Delta Sent_t} \\ & - \underset{(-2.67)}{0.324\Delta Impris_t} + \underset{(1.55)}{1.124\Delta Police_t} + \underset{(1.73)}{0.977\Delta Youths_t} + \underset{(8.05)}{0.626\Delta dum_t} + residual \quad .\end{aligned}$$

(Student-t values in parentheses. Standard Error of Regression = 0.0669).

TABLE 7
Residential Burglary: Short-run economic model 1998-2001
(Plus/Minus Two Standard Errors of Forecast)

	Lower	Median	Upper	Annual Change (%)
1998	439050	503019	576308	-3.1
1999	421689	493414	552078	-1.9
2000	422813	489618	554437	-0.77
2001	418553	485427	550001	-0.86

The predicted modest fall in recorded residential burglaries over the forecast period for this model which excludes the error correction term shows the importance of this term in driving the forecast profile in models where it is included. Purely short-run models of this type have been used by others following the influential work of Field (1990). Such models receive support from several published examples in the empirical literature relating to economic activity and crime which have failed to find the stable long run equilibrium (cointegrating) relationship which is needed to justify the use of error-correction models. Examples of models which failed to find cointegrating relationships include Hale and Sabbagh (1991) for England and Wales and Beki, Zeelenerg and Montfort (1999) for the Netherlands. Scorcu and Cellini (1998) only established stable long run relationships between economic activity and crime for Italy when endogenously determined regime shifts were included in the analysis.

5. Conclusion

Recorded residential burglary offences are subject to quite substantial variations between years, changing in some cases by up to 100,000 offences from a total of 500,000. However, there do seem to be both theoretical and empirical grounds to believe that such changes may be causally related to economic, demographic and criminal justice policy factors such as to make the statistical modelling of this crime a worthwhile exercise. The appropriate way of describing such a relationship is an open question, however. This paper has explored both econometric and time series modelling approaches and developed a number of forecasts or predictions for recorded residential burglary. There appears to be an important difference in forecast levels depending upon whether error-correction models (which incorporate a return to a long-run equilibrium level) to or time series models (which emphasise the inertial aspects of series) are used. For all models, the predictions are associated with wide error bands even when the values of variables used in the forecast periods are treated as being known. The actual uncertainty attached to these variables represents another potential source of error for the predictions. Additionally, all predictions have been made for recorded residential burglary under the 'old rules' of the police for counting offences. The new rules will have some (but unknown) effect in increasing the number of offences, though the effect on this category of crime is thought likely to be relatively small. Despite these qualifications, the next three years should provide extremely useful information of the usefulness of error-correction models in the modelling of recorded crime.

BIBLIOGRAPHY

Beki, C., K. Zeelenberg and K. van Montfort (1999), 'An Analysis of Crime Rates in the Netherlands 1950-93', British Journal of Criminology, Vol. 39, No. 3, pp. 401-415.

Box, G.E.P. and G.M. Jenkins (1970), Time Series Analysis; Forecasting and Control, Holden-Day, San Francisco.

Deadman, D. and D. Pyle (1997), 'Forecasting Recorded Property Crime Using a Time-Series Econometric Model', British Journal of Criminology, Vol 37, No. 3, pp 437-445.

Dhiri, S., S Brand, R Harries and R Price (1999), Modelling and Predicting Property Crime Trends, Home Office Research Study 198, London, HMSO..

Field, S. (1990), Trends in Crime and their Interpretation: A Study of Recorded Crime in Post-war England and Wales', Home Office Research Study 118, London, HMSO.

Field, S. (1999), Trends in Crime Revisited, Home Office Research Study 195, London, HMSO.

GAD (1999). National Population Projections. Government Actuary's Department. Office for National Statistics. Series PP2. No 21.

Hale, C. (1998), 'Crime and the Business Cycle in Post-War Britain Revisited', British Journal of Criminology', Vol. 38, No. 4, pp. 681-698.

Hale, C. and D Sabbagh (1991), 'Testing the Relationship between Unemployment and Crime: A Methodological Comment and Empirical Analysis using Time-Series Data for England and Wales', Journal of Research in Crime and Delinquency, Vol 28, pp 400-17.

HMT (1999). Budget 99. H.M. Treasury. March. H.M.S.O. London.

Macdonald, Z. and D J. Pyle (eds.) (2000), The Economics of Illicit Activity, Ashgate, (forthcoming).

McLeod, G. (1983), Box Jenkins in Practice, Gwilym Jenkins & Partners Ltd, Lancaster.

Mills, T.C. (1990), Time Series Techniques for Economists, Cambridge University Press, Cambridge.

NIESR (1999), National Institute Economic Review, No. 4.

Pudney, S., D Deadman and D Pyle (1997), 'The Effect of Under-Reporting in Statistical Models of Criminal Activity: Estimation of an Error Correction Model with Measurement Error', Discussion Papers in Public Sector Economics No. 97/3, University of Leicester.

Pudney, S., D Deadman and D Pyle (2000), 'The Relationship between Crime, Punishment and Economic Conditions. Is Reliable Inference Possible when Crimes are Under-Recorded?', Journal of the Royal Statistical Society; Series A (forthcoming).

Pyle, D J. and D F Deadman (1994), 'Crime and the Business Cycle in Post-War Britain', British Journal of Criminology, Vol 34, No. 3, pp 339-357.

Scorcu, A.E. and R. Cellini (1998), 'Economic Activity and Crime in the Long Run: An Empirical Investigation on Aggregate Data from Italy, 1951-1994', International Review of Law and Economics, Vol. 18, No. 3, pp. 279-292.

Vandaele, W. (1983), Applied Time Series and Box-Jenkins Models, Academic Press, Orlando, Florida.

APPENDIX

Definitions and Sources of Data of Variables

Residential Burglary: Number of recorded offences of Housebreaking (1950-68) and Burglary and Aggravated Burglary (1969-97: Categories 28 and 29) per capita in England and Wales. Criminal Statistics.

Unemployment: Number registered as unemployed in the UK excluding adult students per capita. Economic Trends

Consumption: UK real personal consumption per capita. Economic Trends.

Conviction Rate: Number of convictions for residential burglary in England and Wales divided by the number of recorded residential burglaries. Criminal Statistics.

Sentence Length: Average length (months) of prison sentence for residential burglary convictions. Criminal Statistics and unpublished data provided by the Home Office.

Prison: Number imprisoned for residential burglary divided by number convicted for residential burglary. Criminal Statistics.

Police: End of Year Strength (excluding special constables). England and Wales. Annual Abstract of Statistics.

Youths: Number of males aged 15-24 years as a proportion of population of England and Wales. Population Trends.

Dummy: Theft Act (1968) dummy. $D = 0$ for $t = 1950-68$.
$D = 1$ for $t > 1968$.