Institute for the Study of Social Change
University College Dublin

Department of Economics,
University College Dublin

An Econometric Analysis of Burglary in Ireland

Kevin Denny (UCD & IFS, London)
Colm Harmon (UCD, CEPR & IZA Bonn)
Reamonn Lydon (University of Warwick)
Ian Walker (University of Warwick & IFS, London)

Preface

This represents the final report in the project "Econometric model of burglary" commissioned by the Department of Justice, Equality and Law Reform.

Dr Kevin Denny is the project leader and is a lecturer in Economics at University College Dublin and a Research Associate of the Institute for Fiscal Studies in London. Dr Colm Harmon is a College Lecturer in Economics, Nuffield Foundation New Career Development Fellow at University College London and an associate of the Centre for Economic Policy Research in London and the Centre for Labour Studies (IZA) in Bonn. Reamonn Lydon is an associate of the Department of Economics at UCD and a doctoral student at the University of Warwick. Ian Walker is Professor of Economics at the University of Warwick. All are associates of the Institute for the Study of Social Change at UCD.

For assistance with data we thank Mr Kieran O Dwyer of the Garda Research Unit, members of the Garda Síochána, and officials of the Department of Justice, Equality and Law Reform. We are also grateful to staff in the Government Publications section of the UCD Library for help with documents. We thank Dr Vincent Hogan and Professor Rodney Thom of the Department of Economics at UCD for econometric advice and Susan Butler of ISSC for administrative support. All opinions and any errors are the responsibility of the authors.

Executive Summary

This report outlines an econometric model of the level of burglary in Ireland between 1952 and 1998. We explain the evolution of the trend in Burglary in terms of (a) demographic factors: in this case the share of young males in the population (b) the macro-economy in the form of real consumer spending per capita and (c) two characteristics of the criminal justice system : the detection rate for these crimes and the size of the prison population. While the share of young males is associated with higher levels of these crimes, the other three variables ((b) and (c)) are associated with lower levels of crime.

Key findings are:

- Macroeconomic conditions as measured by the level of real consumer spending per capita have little effect on the long run level of national burglary crime but, in the short run, growth in consumer spending does depress crime. However consumption has a much more pronounced effect on the average value of a burglary act. Hence the evidence suggests that a favourable macroeconomic environment affects the nature of the criminal activity (i.e. more valuable burglary) but not the overall frequency of property crime.

- Demographic conditions are measured by the share of young males in the adult population and this variable has a positive effect on crime both in the long run and the short run.

- The criminal justice system impacts on crime through two channels – the detection rate (reflecting the probability of being caught), and the numbers in custody for crimes against property with violence (reflecting the consequences of being apprehended). Both have well determined negative effects on crime reflecting the deterrence effect and the direct effect of incarceration.

- Some difference can be noted when we consider the data separately for rural and urban divisions. Detection has twice as big an effect on crime in the rural areas possibly reflecting the more risk averse nature of criminals in these areas.

- One somewhat surprising result is that we were unable to find any robust effect from direct measures of labour market activity such as unemployment rates or wage levels.

1. How do economists think about crime? A Brief Review

1.1 Introduction

Crime matters to people. The direct victims of crime, in addition to the immediate impact of the violence, loss of property and so on, report levels of anxiety and stress for some time after the event. The effects of the crime are not limited to these victims (and their families) alone. Society as a whole feels threatened by the level of crime. This can affect people's lives in many ways. Individuals may be reluctant to venture out for fear of assault or fear that their property will be stolen. They may go to extra expense to minimize the chances of being victimized. There may also be a lower level of trust in the community: to use current terminology "social capital" is damaged. Those responsible for crime are not drawn randomly from the population. The international evidence is that being young, male, with low education and low income is associated with being involved in crime. So wider considerations of social justice also suggest that crime is an important policy concern.

However rising crime is not inexorable. Although it is currently high by historic standards it has fallen significantly in recent years. The actions of the courts, the probation services, the Garda Síochána, the legal system are likely to matter. There are other factors, amenable to public policy, which may also influence crime such as the level of unemployment, education or inequality.

It follows then that understanding the determinants of the level of crime is important in thinking about reforms to public policy to further reduce crime levels. We emphasize that policy towards criminal justice is a complicated matter involving a host of legal, social and other issues. The causes of the level of one particular type of crime constitutes one piece of the jigsaw puzzle. Perhaps surprisingly, research on the determinants of crime is in its infancy in Ireland. In this report we carry out one of the first econometric analyses of crime in Ireland and the first for nearly twenty years. By *econometric* we mean using the modern techniques of multivariate statistical analysis as applied to economic models. This sort of analysis has been carried out many times by economists and criminologists for many other countries.

It is worth explaining briefly what we mean by an economic model and why this is the "lens" through which we view the data. Economic models see individuals as essentially rational who make decisions on the basis of what they perceive the benefits and costs to be of their actions. The major breakthrough was the work of Gary Becker, a Professor of Economics and Sociology at the University of Chicago and now a Nobel laureate in economics. In Becker's model, criminals are rational individuals acting in their own self interest. In deciding to commit a crime, criminals weigh the expected costs against the expected benefits. So some individuals will choose to commit crimes because they value the benefits – say property taken - more than the expected cost to them. The cost of crime will be an *opportunity cost*, it is what they give up to pursue the crime. This can include working in the legitimate economy. If jobs are easier to come by and/or such jobs are well paid then the opportunity cost of crime is higher.

However crime is essentially a risky activity for the criminal and he/she needs to take into account the chance of being caught and the outcome for him/her in the event of being caught such as imprisonment. Of course there are other factors which will influence individuals decisions to commit crime, the attitudes of society to crime, the attitudes of one's family and so on. However we do not explicitly consider them here either because there is simply no way of measuring them or because they are intrinsically "micro" level variables which will influence individuals but are unlikely to explain trends in the aggregate level of crime. Here we are concerned with burglary and crimes against property more generally and it seems plausible that the "cost/benefit" approach is a sensible way of thinking about it. It is clear that there are types of crime, sexual or political crimes for example, for which this would not be the most useful framework.

A different perspective on understanding crime tends to emerge from sociological analyses which sees crime as a form of social dysfunction. This idea is associated with the French sociologist Durkheim's notion of *anomie* - feelings of alienation and rootlessness amongst individuals. There is no necessary contradiction between such a view and the approach taken here since it is difficult if not impossible to test such a model. For example it is difficult to know how one could consistently measure the trend in *anomie* over a period of 30 or 40 years. Possible indicators would be the level of suicide or the level of inequality, the former a consequence and the latter a cause of the phenomenon. In the first case however we know that measured trends in suicide are probably a poor guide to the underlying phenomenon since they will reflect changing attitudes, changes in reporting practice and so on. In the second case, although both economic and sociological analyses would see a rôle for economic inequality affecting crime there is simply not enough data to use.

An important aspect of the present study is that it is *multivariate*, that is it seeks to model crime as depending on a series of variables. So when we measure the effect of one variable it is holding other variables constant. This is important as simply finding a correlation between two variables is no guide to whether there is any relationship between the two. In the social sciences it is essential to avoid such naïve comparisons as they are generally misleading and unreliable.

1.2 Evidence on the Economics of Crime

As discussed, the "cost" of crime to criminals consists of two parts. One is the income foregone by devoting time to criminal activity (the opportunity cost). The second cost is the time criminals expect to be incarcerated because of their activity. "Expected punishment" is not the same as the length of time a convicted criminal actually spends in prison. Most crimes never result in an arrest and many of those arrested are not prosecuted. Many convicts receive non-custodial sentences.

Expected punishment, from the criminal's viewpoint, is a probability, not a certainty. For example, in figures cited by Rubinstein and Woodson (1995) only 7% of U.S. burglaries result in an arrest according to the National Center for Policy Analysis (NCPA). Of those arrested, 87% are prosecuted and of those prosecuted, 79% are convicted. Of those convicted, a mere 25% are sent to prison (most are paroled). After multiplying these probabilities, we see that a potential burglar faces only a 1.2% chance of going to prison for each act of burglary committed and that once in prison, she/he will stay there for about 13 months. But since she/he will escape imprisonment more than 98% of the time, the expected "cost" of each burglary to the burglar is only 4.8 days.

The rational criminal will ask him/herself whether an act of burglary is likely to net him/her goods worth more than 4.8 days behind bars. If the answer is yes, then the crime pays. The goal of the criminal justice system is to raise expected costs of crime to criminals above the expected benefits. People will commit crimes only so long as they are willing to pay the prices society "charges."

Since Becker's seminal work (published in 1968) economic research into the topic of crime has focused on a simple model with three elements – the supply of criminal activity, the size of the punishment and the probability of arrest and conviction with the latter two elements forming the main weapons available for deterrence of crime. Supply of crime should, in this simple economic model, decline if the offence is more likely to be detected or more likely to carry a heavy sentence hence the motivation toward policies such as increased police coverage, closed circuit television and reviews of sentencing strategies.

However the market will have its imperfections. Criminals for example may be poorly informed about the chance of being caught and the implications of this for their liberty. Moreover for criminals to react in the manner of the simple model suggest that they are risk averse but this may not be the case. Individuals who enjoy risk may be attracted to crime as a lifestyle.

If criminals are rational they will also respond to the relative 'prices' of different crimes and move from crimes that are heavily punished to lighter crimes in punishment terms. Levitt (1995) found for the US that deterrence is more important than incapacitation for minor crimes and the deterrence effects are generally negative. Deterrence effects were large for burglary (each arrest is estimated to eliminate two burglary crimes).

1.2.1 Irish Research

Criminology is a very under-developed field of research in Ireland and the rigorous statistical or economic analysis of data is the exception rather than the rule. The nature of the activity, being illegal, necessarily makes data collection difficult and hampers micro-level research in particular. There are nonetheless several important contributions worth noting.

Virtually the only economic analysis of crime in Ireland that we are aware of is Bacon and O' Donoghue (1977). This applies the model of Becker (1968) to Ireland to analyse what the optimal levels of expenditure on crime control should be. While most of this paper is not directly relevant to the proposal, they do include an econometric analysis of the determinants of crime, distinguishing between violent and non-violent crime. They find that unemployment has positive and negative effects respectively on these two categories. This is somewhat at odds with results for other countries though the crime categories do not correspond exactly.

The ESRI report by David Rottman (1980) analyses the aggregate trends in the data with a largely sociological focus. It pays particular attention to the quality of the data and notes the relatively good data on crime that can be derived from the annual reports of the Garda Síochána. It also carries out a basic statistical inquiry into the trends in the data. In particular it regresses the level of crime in nine different categories on the level of unemployment.

For most types of property crimes in particular there is a well determined and positive effect. Interestingly for assault, he finds a negative impact, a finding which is mirrored in subsequent research on US data using much more advanced estimation methods (see for example Raphael and Winter-Ebmer 1999). The techniques used by Rottman are, certainly to a contemporary reader, relatively unsophisticated. Nonetheless it provides a good introduction both to the data and the possibilities for statistical analysis.

McCullagh (1992) discusses the relationship between imprisonment and unemployment between 1951 and 1988 and concludes that there is a positive relationship but only in the latter part of the period. Disappointingly, his method is based on a visual inspection of the data and completely eschews any statistical testing. The calculation of simple correlation statistics would have been very useful and is easily done. His conclusion that "analyses based on forms of multi-regression may disguise more than they reveal about the data under examination" is not well founded. Since one does not expect crime to be determined solely by any one variable, the simple correlation between crime and unemployment, say, is of no particular significance.

An interesting micro level study is the paper by Bacik *et al* (1998). They modelled the probability of a custodial sentence as a function of characteristics of the individuals and the communities from which they are drawn. The authors collected a large sample (over 1000) of cases from the Dublin District Court. Their results show that those who come before the courts are not typical of the population, being predominantly male, young and from more deprived areas. While this by itself is unsurprising, as the authors note, the degree to which this holds is striking. Using multivariate methods (logistic regression, a technique used where the variable to be explained is binary) they show, *inter alia*, that among those appearing in court, being from a deprived area is associated with a higher probability of receiving a custodial sentence. This probability is *decreasing* in age but is higher if the crime in question was a property offence.

While the results are very interesting, interpretation is less straightforward. For example the age result could reflect a judges' reluctance to incarcerate older individuals perhaps because of family responsibilities. It could equally reflect the nature of the crime, maybe younger people are associated with more serious offences. The analysis includes controls for class of offence (property, drugs, public order etc) but *within* these classes there will be variation in the degree of seriousness. Another problem is that in this data not receiving a custodial sentence includes those convicted but not imprisoned and those who were acquitted.

Nonetheless this study provides an excellent example of the ability of modern statistical methods to measure the influence of different factors on a variable of interest. It would be highly desirable if micro-level data were collected to allow further research along these lines.

Also worth noting is a recent study by O'Mahoney (1997) which studies the characteristics of a sample of prisoners in Mountjoy Prison. This emphasises the fact that at the micro level, crime is strongly associated with individuals who are young, poorly educated and generally economically disadvantaged.

The significance of these last two micro-studies for the proposed research is that they largely confirm the finding of much international research. As the Whittaker (1985) report puts it "most crime at present originates amongst unemployed youth in disadvantaged areas".

Brewer *et al* (1999) is a historical/sociological monograph on crime in Ireland, much of it on Northern Ireland. However it does contain, as well as much useful background information, an extensive Statistical Appendix which if combined with Rottman's study provides a good source of aggregate data on crime. In this report we have collected all data from their original published sources

1.2.2 Crime and the Labour Market

The relationship between crime and the labour market has been a major issue in the US and UK research. Freeman (2000) suggests that there is little *direct* evidence linking education to reductions in crime and the perceived linkage relates to the effect that education has on factors such as unemployment and inequality. There has been a dramatic rise in crime over the period 1950 to 1997 – reported crime rose by a factor of 3 in the US and by a factor of 2 in the UK. Moreover the significant rise in wage inequality that has been observed in the US and Great Britain over the past 25 years suggests that the return to legitimate work has fallen for low or unskilled individuals. This is especially true for men.

Estimates of the return to crime are harder to gauge – again US figures suggests that criminal earnings for inner city youths rose by an average of 5% over the period of the 1980's. This, when coupled with downward changes in the probability of incarceration for youths in the US during the same periods suggests that crime rates do react to labour market condition and that this behavior may indeed be rational for some individuals. Upward trends in inequality are associated with higher levels of both property and violent crime (see Kelly, 2000). Winter-Ebmer and Raphael (1999) find positive effects of unemployment on crime that are not just statistically significant but large in size. Leigh (1998), in a review of work published in this area, concludes that increased education is positively and strongly correlated with absence of violent crime, measures of health, family stability and environmental benefits.

Lochner (1999) develops and estimates a model of the decisions to work, to become educated and to commit crime and allows for the possibility of interactions between all of these choices. The model suggests that education is correlated with crimes *that require less skill*. Part of the model allows for simulation of the effects of education subsidies on external outcomes and predicts that education subsidies reduce crime. Insofar as possible, empirical implications were explored using various large scale US micro datasets. Ability and high school graduation significantly reduce the participation of young men in crime and the probability of incarceration. Evidence from the census data supports a general finding that states with higher rates of high school participation and tougher penalties have the lowest index for property crime.

1.2.3 Does Crime Respond to Punishment and Policing?

There is evidence that crime responds to punishment but the major issue here is that more crime is induced by low incarceration rates but higher levels of crime may in turn lead to higher incarceration rates as some of the criminals are convicted. This problem, known in the econometric literature as simultaneity, means that the existing evidence on crime and punishment may be biased. It is a special case of the wider statistical problem of measuring the effect of a "treatment" when it is applied non-randomly.

There are statistical techniques that can sometimes be used to solve this problem. For example if one can find a variable that independently influences imprisonment rates (independent of the crime rate that is) then one may be able to use this to eliminate the simultaneity bias using a procedure known as Instrumental Variables estimation.

Levitt (1996) suggests using exogenous factors that limit the use of incarceration in the United States such as caps on prison populations as providing an independent influence on incarceration. His study uses the number of state litigation suits for violation of prison directives as a means of isolating the effect of punishment levels on crime rates. In this work a 10% increase in imprisonment rates would lead to a 0.7% decrease in property crime and a 1% decrease in violent crime using conventional methods. However using the more sophisticated methodology suggests that these figures are underestimated with the 'corrected' effects of a 10% increase in imprisonment rates being 4% for violent crime and a reduction of 3% for property crime. This shows the importance of paying careful attention to the use of appropriate statistical methods since they lead to estimates of the effects which are a multiple of the naïve estimates.

The same issue of simultaneity applies to the relationship between the effects of policing on crime. More police is likely to have a negative effect on the number of crimes however there may be a simultaneity problem if the authorities respond to a crime surge by increasing levels of policing thus generating a positive correlation. In addition, a higher level of policing may encourage the public to report crimes more so that crime statistics may rise due to lower under-reporting. Again, the same sort of statistical procedure can be used if one can identify influences on police levels that are not directly related to crime levels.

In his 1997 paper, Levitt suggests that during election years police forces tend to swell. He therefore uses US local elections to model the problem and again finds large differences between the conventional models and the ‘corrected’ one. For property crime he estimates that a 10% increase in the number of police officers reduces property crime by about 1.5%-3.8% depending on the precise methodology used. This implies a decline in reported property crimes per police officer of between 4 and 12.4 per officer annually. Taken in conjunction with similar effects on violent and other crime that Levitt estimates in this study, this all translates into an additional social benefit of over $200,000 from the hiring of one additional police officer[1]. The extent of the problem (and the applicability of this solution) may be less in Ireland where it takes some time to increase the size of the Garda Síochána.

A number of other experiments have been carried out largely in the United States. For example California’s ‘Proposition 8’ imposed enhancements to sentences for a select group of crimes particularly for re-offenders. In the year following the passage of the legislation the Proposition 8 offences recorded reductions of 10% relative to unchanged crimes. Three years after the law came into effect eligible crimes had fallen by 20-40% compared to other crimes. This brings out a significant advantage that researchers have with de-centralised judicial systems, changes in state laws act as "natural experiments" allowing one to observe the effects of changes in variables in one state compared to another. This is much more difficult to do where reforms are only at a national level as in Ireland.

1.2.4 Drug Use

There has long been in interest amongst researchers between possible links between drug use and crime. Corman and Mocan (2000) investigated relationships between crime, drug use, police and arrests in New York using monthly data from 1970 to1990.

[1] Levitt also makes the point that his study is based on reported crimes only. As criminals do not know if a crime will be reported or not it seems logical to translate the effects on reported crime into unreported crime. In any event reporting bias, if present, will understate the true effectiveness of police in reducing crime.

They found that drug usage increased from the mid-1980's to about 1988 and at the same time drug arrests increased substantially. During the same period there were increases in murders, assaults and motor theft. Arrest rates for other crimes did not decline during the period of increased drug arrests.

In a multivariate analysis they found, holding constant arrest rates and policing levels, property crimes increased when there were unexpected increases in drug usage but this was not the case for murders or assaults. They also found evidence of police deterrence effects on property crime only. Thus, in this time series model a causal relationship between drug usage and property related crime is noted.

2. Irish Crime Data

2.1 Introduction

In this section we describe the availability and quality of data that is either necessary for an exercise such as this or that might be conceivably useful given the issues arising in the first half of the report.

We first discuss in section 2.2 data specifically related to crime including alternative measures of our dependent variable and then go on to discuss the situation vis-à-vis the demographic and economic variables that are likely to be used as covariates or explanatory factors. As with most such studies the amount of data declines as one goes back in time so there is an inevitable trade off: if one wishes to focus on more subtle measures, say of economic activity, then one may be constrained to use a much shorter time-span. A further issue concerned the quality of the data since even if the data exists one may have reason to doubt the reliability of the information in which case there may be an argument for foregoing its use.

Although the Central Statistics Office was founded in 1949 economic data in Ireland is relatively scarce before 1960.[2] In our search for data we have taken 1950 as our starting point since there is very little before this. Econometric analysis of the Irish economy typically starts with data from around 1960 or later it is unlikely that trying generate a dataset that starts before 1950 would be practical. Perhaps ironically, consistent series on crime variables precedes many of the key macroeconomic time series.

2.2 Crime data

The source of the crime data is the annual Reports of the Garda Síochána. Crime data is also published in the Statistical Abstract of the CSO but this source offers no particular advantage. These reports are in general a very useful consistent source of the basic data which we wish to analyse. Note that there is considerable debate in the international research literature as well as amongst Irish researchers about the reliability of official statistics. Under-reporting is well known to be a potential problem but to an extent to which, almost by definition, is hard to ascertain. Here we have taken the data at face value since there seems little alternative and evaluation of the data is beyond the remit of the study. Further research on this issue is clearly desirable, possibly though regular surveys such as the recent ESRI survey, Watson (2000). It may also be useful to analyse data on insurance claims as an additional source of information. The only study that we are aware that explicitly addresses the question of statistical inference associated with under-reporting of crime is Pudney *et al* (2000). Their investigation, using British data, leads them to the conclusion that it is "of little practical significance".

2.2.1 Burglary data

The Garda reports classify crimes into four groups of which the second, Group 2, Crimes against property with violence, is relevant for us. We have extracted the data for burglary from these reports from 1950 until 1999. However we propose to use only data up to and including 1998 for two reasons. Firstly the data for 1999 refers only to the period January-September due to the introduction of the PULSE system of collecting crime data. Clearly it is not directly comparable with data based on a full year given especially that crime displays a strong seasonal component, being greater in the winter. We hope that this data will subsequently be revised at some point to refer to the whole year.

In principle one could attempt to seasonally adjust the data for 1999: by using within-year variation in crime for earlier periods to impute the "missing crime" for the last quarter of 1999. However this would be a

[2] Prior to this much of the collection and publication of data fell to the Statistics Division of the Department of Industry and Commerce.

significant undertaking for a relatively small return and one that would introduce an additional source of error into the data. The second reason is that not all our covariates can be observed for 1999.

There is one issue in the construction of the basic burglary series data worth noting namely that there is a change in reporting in 1977 which amalgamated a number of different series to create a new burglary series. Up to and including 1976 there is data on the following offences: (i) Sacrilege (i.e. theft from places of worship), (ii) Burglary, (iii) Housebreaking, (iv) Breaking into shops, warehouses or other premises, (v) Attempts to break into shops and houses, (vi) Entering with intent to commit a felony and (vii) Possession of House-breaking tools. From 1977 onwards there is one series which, as far we can see, is the sum of the first five of these. Of the six categories listed above (iii) and (iv) account for the bulk of the total. We give data for the relevant period below

Table 2.1 Monitoring the Definition Changes in Burglary Data

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	Sum of (i) to (vii)
1975	76	556	5840	7194	348	1116	30	15160
1976	57	10291	4871	3372	168	382	36	19177
1977		21009					22	21031
1978		18923					22	18945

There is one further complication to be aware of, due to the changing reference period used in Garda Reports. Up to and including 1957 the reporting year ended in December. From 1958 to 1974 inclusive the reporting year ended in September after which it reverted to the end of December again.

2.2.2 Regional breakdown

We were asked by the Department of Justice, Equality and Law Reform if it was possible to generate some form of regional or spatial breakdown of the data. This could be an important issue since one might expect differences in the nature or urban and rural crime. However this is not straightforward to do. Burglary data is only disaggregated spatially from a relatively late date which would give too short a period of data to analyse econometrically. However an alternative approach is possible. As noted above Burglary comes under the class of data recorded as Offences against property with violence or "Group 2" in the statistics. These series are consistently provided at a Division level and we have collected them for Dublin, Cork, Limerick, Galway and Limerick. We have aggregated these five series under the heading “Urban”. By subtracting the “Urban” figure from the totals for Ireland we get a “Rural” series. We think this is the best feasible way in which a spatial disaggregation can be pursued. **Figure 2.1** graphs the time series of both Burglary (based on our amalgamation of the pre-1977 data discussed above) and the Group 2 data from 1950 on. The share of the former in the latter is not constant see **Figure 2.2** and drops significantly in recent decades.

Figure 2.1 - Time Series of Crime Data

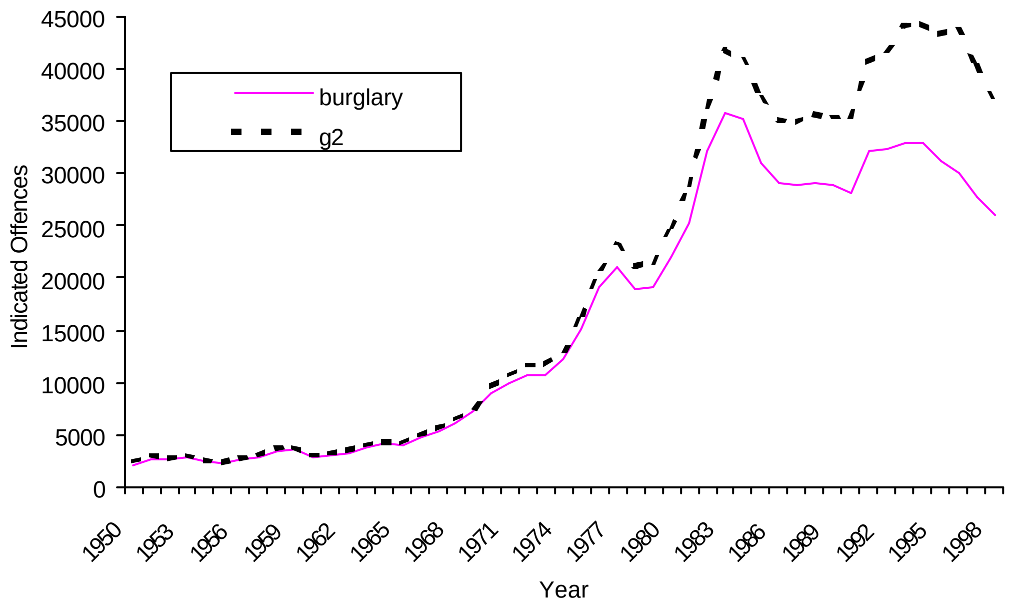
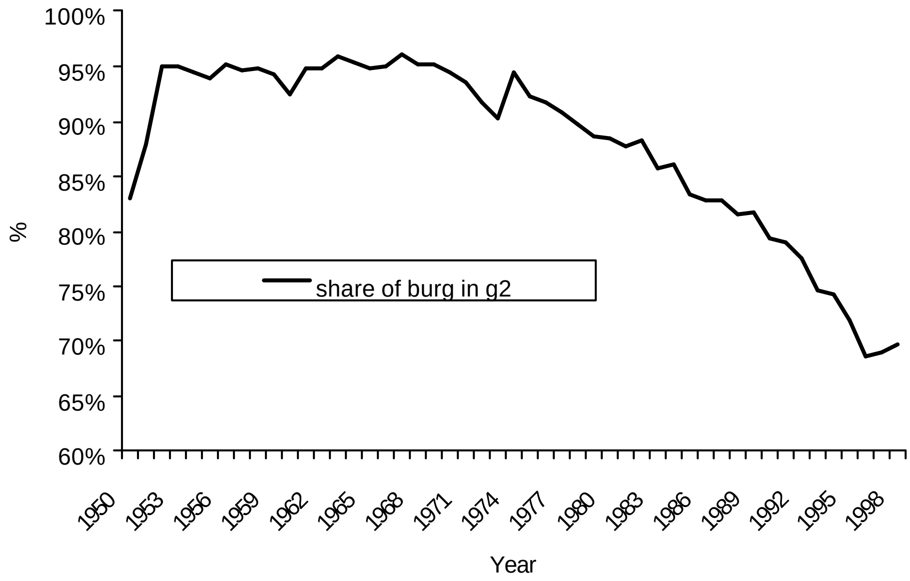


Figure 2.2 – Burglary as a percentage of Crimes against Property with violence (g2)



The fall in both series from the mid 1990s is evident. Unpublished data from the Department of Justice, Equality and Law Reform for 1999 and 2000 indicate that this trend is continuing, with the numbers of burglaries for those years being 23,042 and 20,477 respectively. So burglary for last year (2000) is now down to 79% of the 1998 figure and is lower than it has been since 1982.

To give an idea of the differences between burglary and the category of crimes against property with violence **Table 2.2** below gives a breakdown for two years. It can be seen that burglary accounts for the bulk of these crimes, over 88% in 1980 and a small number of other categories account for the remainder. As can be seen in Figure 2.2 the share of burglary in crimes against property falls over the 1980s and 1990s and this would appear to be largely accounted for by a greater incidence of malicious damage against property (including schools). It needs to be borne in mind therefore that in analysing the trend in crimes against property one is looking at a class of crime whose internal composition is to some extent changing. However the reasons as to why malicious damage to property increases in significance is one that is beyond the scope of this study.

Table 2.2 Crimes against property with violence ("Group 2")

	1980		1998	
Burglary	2,1974	88.3%	25,730	69.2%
Aggravated burglary	201	0.8%	657	1.8%
Robbery	939	3.8%	1,831	4.9%
Malicious damage to property	1,318	5.3%	8,223	22.1%
Other	446	1.8%	750	2.0%
Total	24878	100%	37191	100%

A convention in much of the literature is to normalise the burglary series by dividing by the population and this is shown in **Figure 2.3** while **Figure 2.4** shows the share of "Urban" (as defined above) in total group 2 crimes.

Figure 2.3 – Burglary per 1000 Population

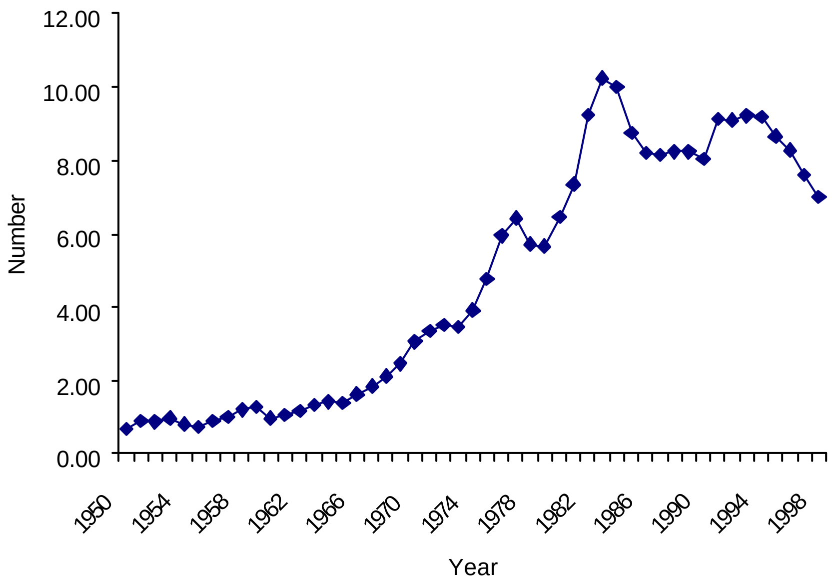
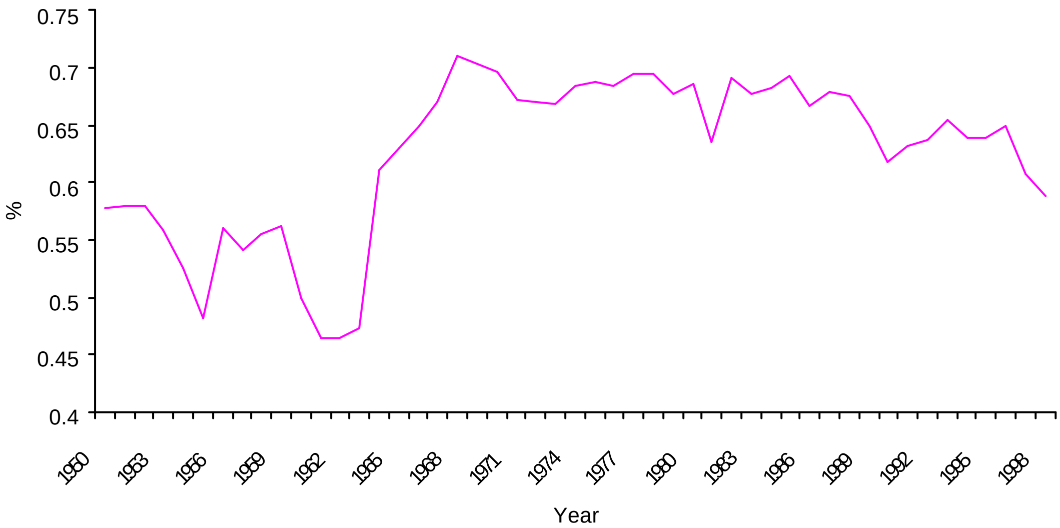


Figure 2.4 – Share of "Urban" in Total crimes against property with violence



2.2.3 Value data

The simple frequency of crimes, that is the number of offences recorded, can be misleading or at least may be only partially informative, since it says nothing about the severity of the crime. An alternative approach is therefore to have a measure of crime that weights individual offences in some way with “bigger” robberies contributing more to the overall trend than lesser crimes. For crime against property a natural weighting would be based on the value of property stolen so that a crime in which £100 is taken is equivalent to two crimes in which £50 each is taken. This is not to say that this is the only, or even the ideal, measure of the trend in the crime in question. If it is the experience of being burgled (or otherwise victimized) rather than the property stolen that is of concern to individuals and society then the simple frequency of crimes may be the appropriate measure. Alternatively one might wish to weight the crimes by the numbers of victims so that the burglary of a household with two people counts for more than that of a single person household. This however raises the question of how one should weight non-residential properties.

So underlying the decision of whether to use a weighted or unweighted measure is to some extent a value judgement and also a question of data availability - it is not possible for example to weight burglary crimes by the numbers of residents in the household. Our view is that both approaches have merit and hence we also estimate the same model specification but with a weighted measure - the average value of property stolen- as dependent variable.

The Garda reports contain a series on the value of property stolen for Group 2 crimes though not for burglary. Data is also presented on the number of crimes on which the value series is based – these are a subset of all Group 2 offences. Data is also available on the value of property recovered. This may be useful as an index of deterrence since there may be less incentive to steal if there is a high probability of it being recovered.

Figure 2.5 – Value of Stolen Goods (Deflated by the Consumer Price Index, 1950=1)

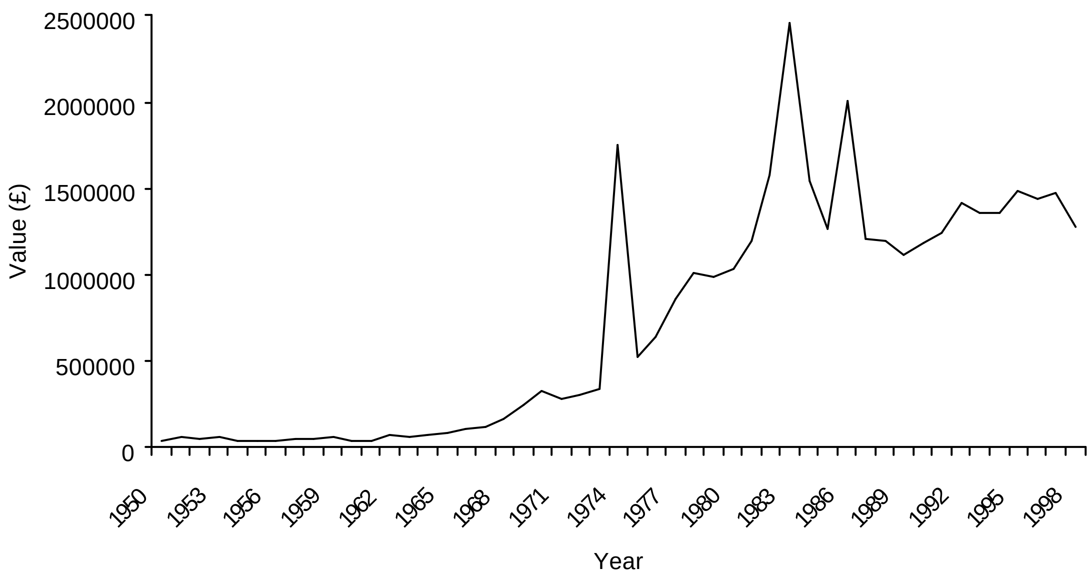


Figure 2.6 – Average Value of Crime Per Reported Case

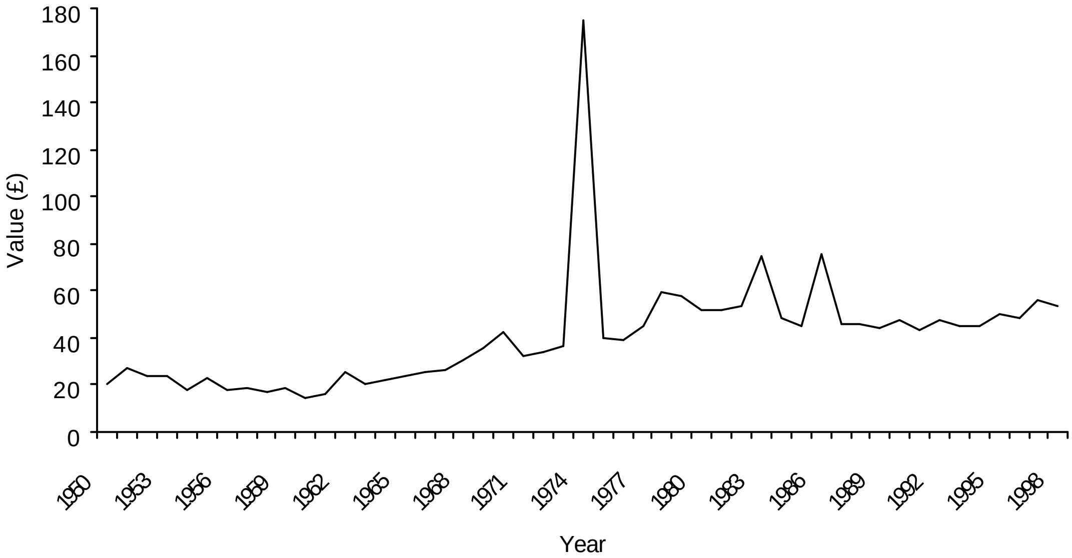


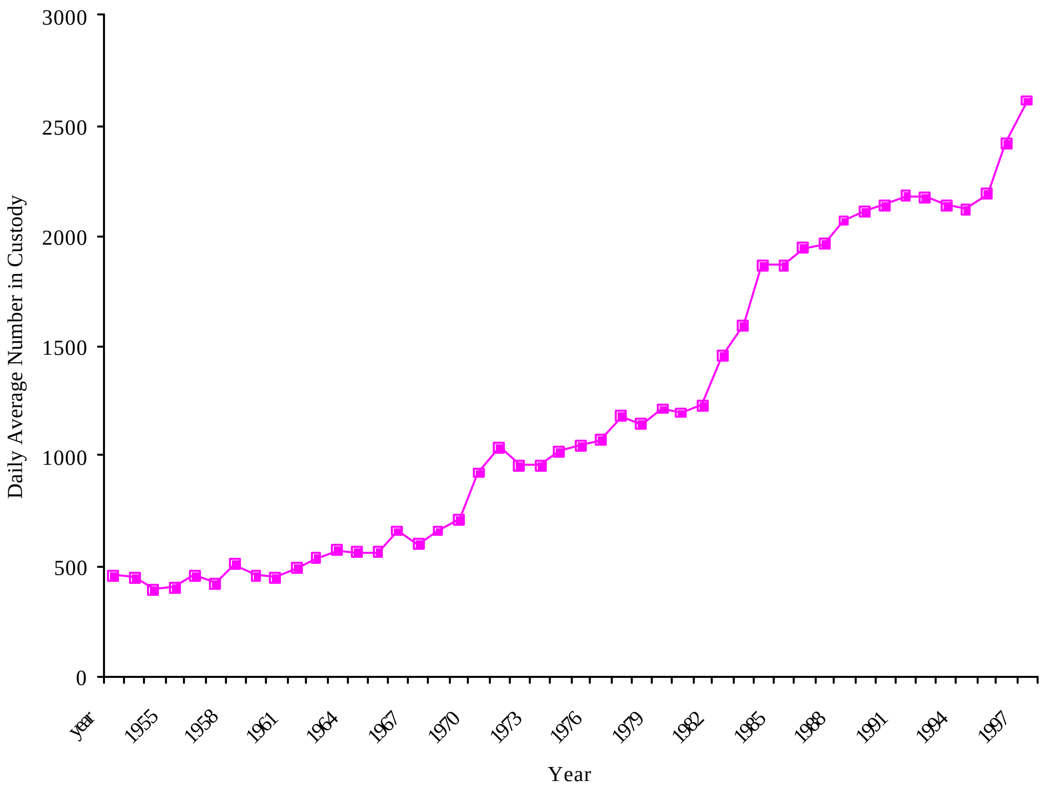
Figure 2.5 graphs the value series, deflated by the Consumer price Index (indexed 1950=1). The three “spikes” in the data at 1974, 1983 and 1986 are striking. An obvious question which can be asked is whether and to what extent these spikes can be explained by an increased frequency of crimes. So in **Figure 2.6** we graph the value series divided by the number of cases, which hence corresponds to the average value of property stolen (in real terms). This to some accounts for the spikes in 1983 and 1986 which are now much less pronounced. The Garda reports for 1974 and 1986 identify one particular crime in each of those years and give the value taken (see data appendix) so we simply adjust the series by subtracting the amount. For 1983 we were unable to explain the spike, instead we replace the observation for that year with the mean of the 1982 and 1984 data. More complicated methods of imputation are available but we believe this method is adequate.

2.3 Explanatory variables

As explanatory variables we propose to consider a small number of key demographic and economic variables which the international literature has suggested as determinants of property crime. In addition one may wish to consider variables representing the criminal justice system which may act as deterrents to criminal activity. We use two variables under this heading one relating to detection of crime and one to imprisonment.

For the latter we use the daily average number of people in prison from the CSO's Statistical Abstract augmented by the report of the Prison Service. **Figure 2.7** graphs this for the period of interest.

Figure 2.7: Daily average numbers of people in custody

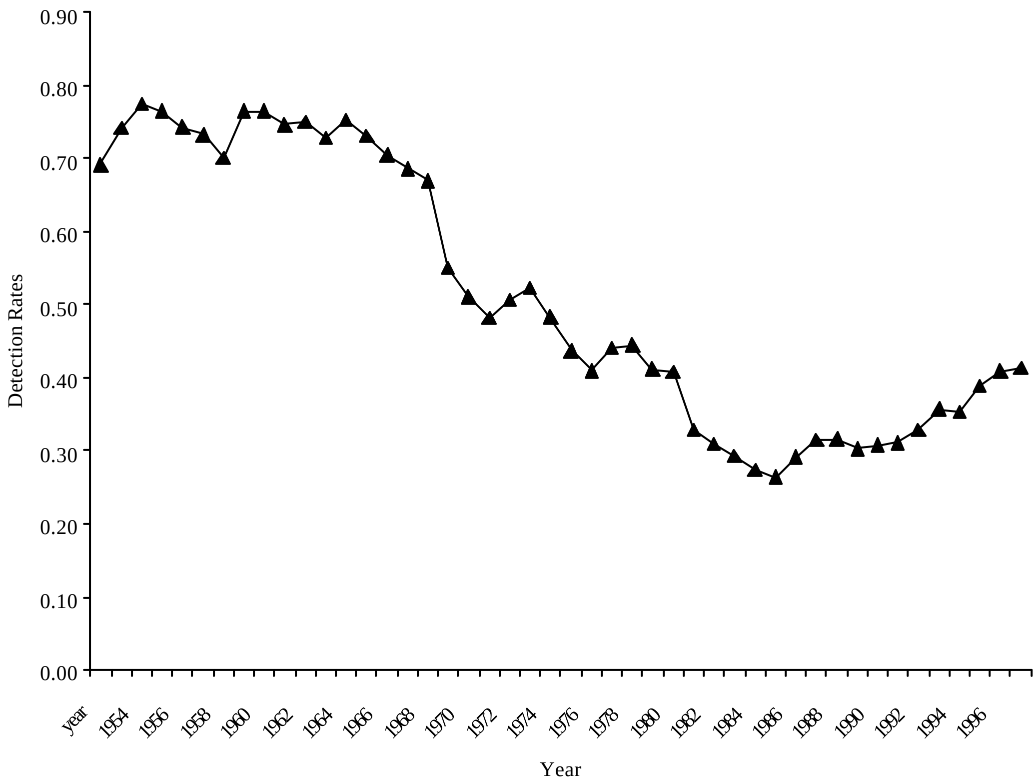


9

7

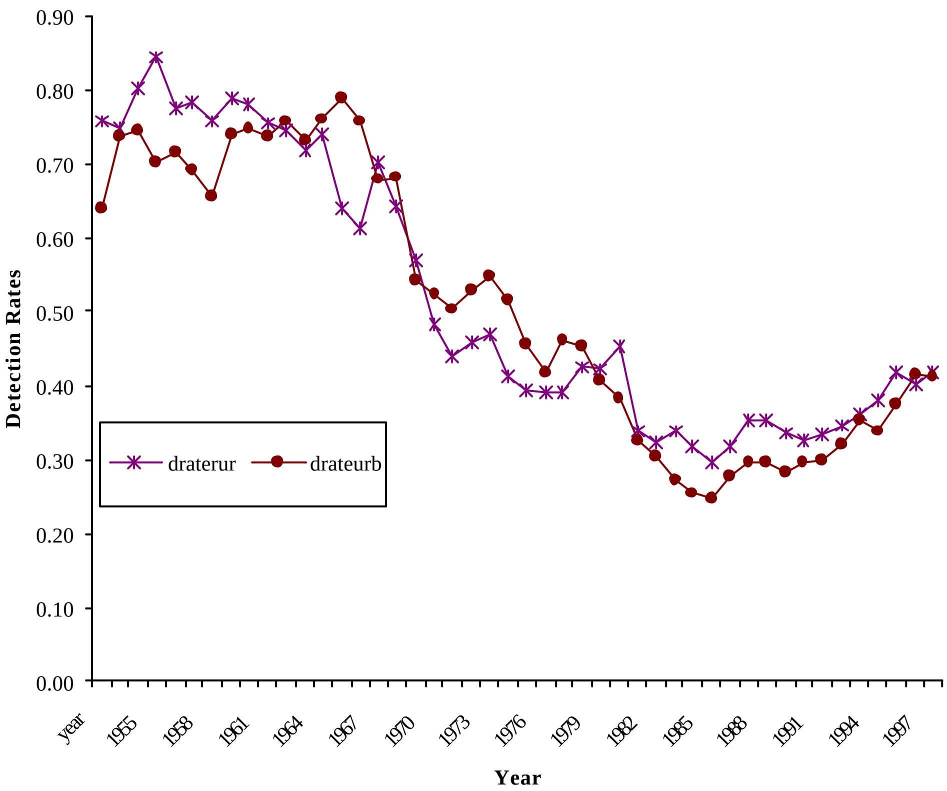
For detection rates we use the number of indictable offences in group 2 detected by the Gardaí, there is no figure available specific to burglary. The data can also be broken down regionally which is useful. The detection data as a proportion of the corresponding crime level is given in **Figure 2.8**:

Figure 2.8 Detection rate for crimes against property with violence



The detection rates broken down into an Urban series ("drateurb") and a Rural series ("draterur") earlier are in **Figure 2.9**. As can be seen they follow a very similar pattern.

Figure 2.9 Detection rate for crimes against property with violence : rural and urban

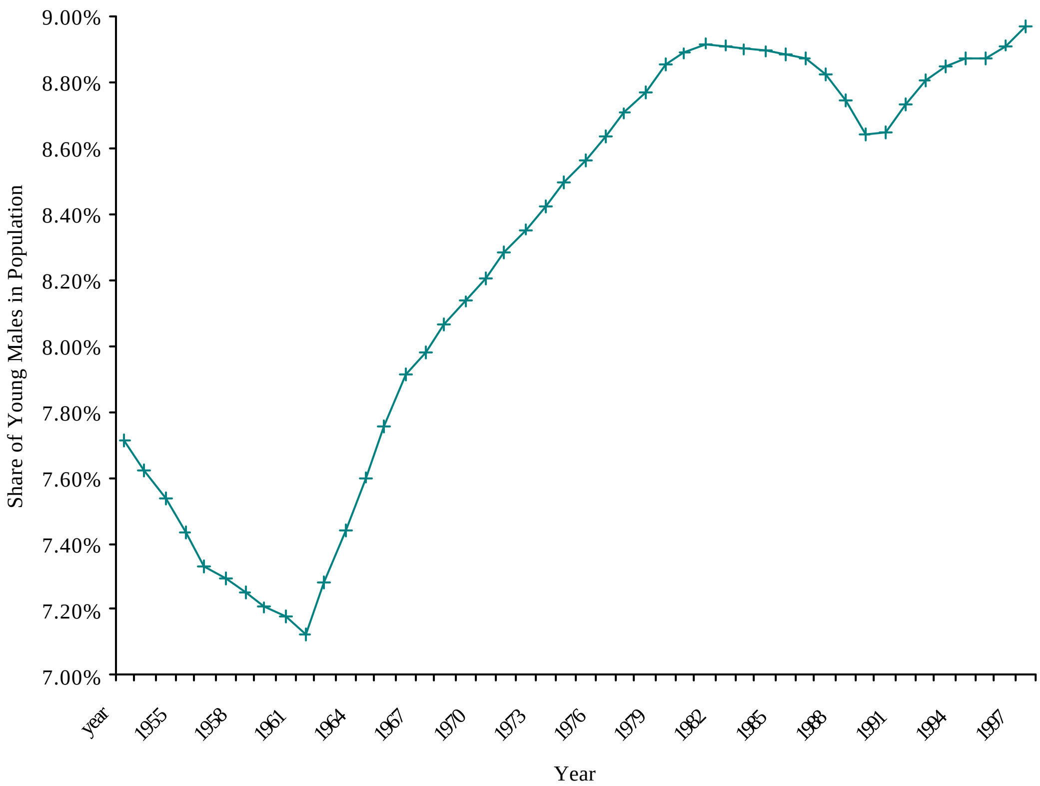


2.3.2 Demographics

There are a number of reasons why one might want to use demographic data. As noted above one may want data on total population to scale the number of crimes so that one analyses "burglary per head of population". Whether to model this or the simple level of burglary is to some extent a matter of taste and the literature is divided on which is the best to use. In this study we do *not* divide the crime data by population. A more important reason is that the age and gender composition of the population is likely to be important since the international evidence is that crime, at least of the sort considered here, is largely committed by young males. The evidence for Ireland is consistent with this. Fortunately all, or almost all, of the demographic data that one needs is available in one source the ESRI Time Series Databank, developed and provided by John Fitzgerald and Jonathan Hore. This gives population broken down into four age categories (0 to 14, 15-24, 25 to 64, 65+) for males and females separately from 1950 to 1999.

We construct a variable which is the share of young males in total population. “Young” in this context means between the ages of 15 and 24 inclusive. This series is graphed in **Figure 2.10**.

Figure 2.10 Share of young males in total population

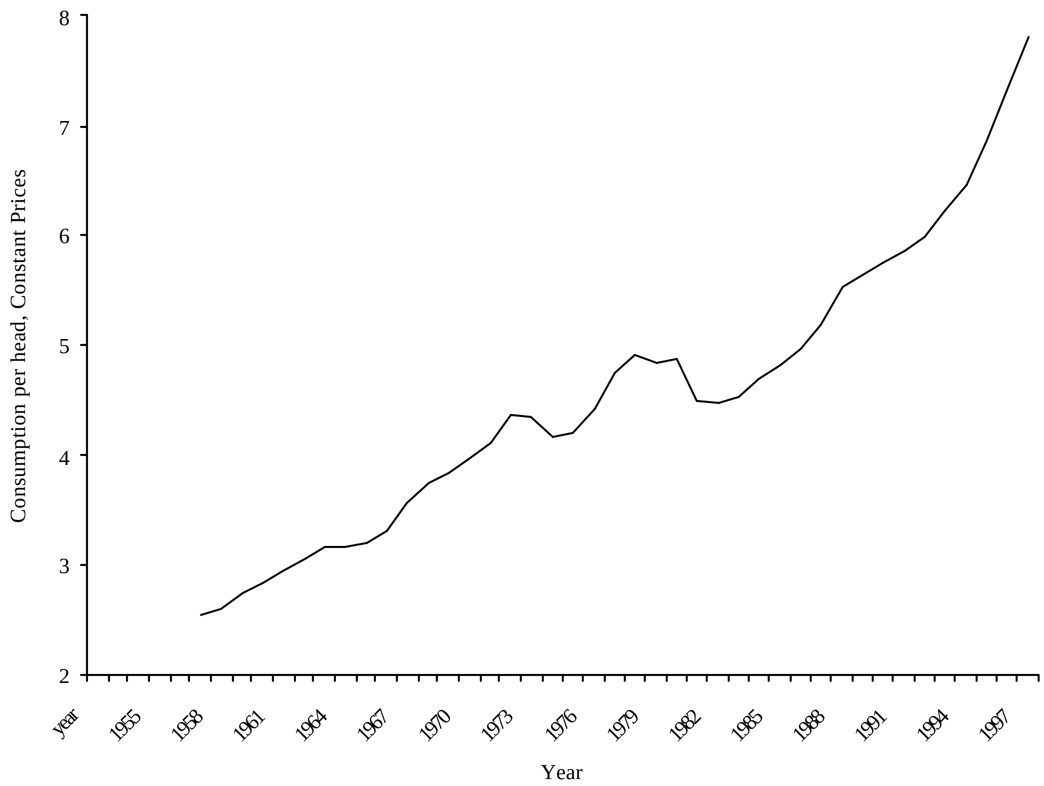


2.3.3 Economic activity variables

2.3.3.1 Consumption expenditure

We have experimented with a number of macroeconomic variables. Perhaps surprisingly only one was found to feature consistently so we end up with quite a parsimonious specification. Amongst the variables we considered are the level of real wages, the rate of unemployment and the level of migration. The variable that we did use is consumer expenditure, in real terms, per capita. It should be noted that the recent United Kingdom Home Office studies (Field(1999), Dhiri *et al*(1999)) also find a crucial role for consumer spending. Another variable which would be a close substitute for consumer spending would be GNP per capita. **Figure 2.11** illustrates the series that we have used.

Figure 2.11 Real per capita consumer spending



3. Econometric Analysis

3.1 Overview of empirical modelling

In this section we present the econometric analysis that is the core of this report. We first outline some important statistical preliminaries. We then present the estimated model of burglary using data from 1957 to 1998. As will be clear from the section on methodology this proceeds in two steps, the first estimating the long run relationship between a set of variables and the second considering short run responses. To bring out the implications of the results we then use the estimates to simulate a number of scenarios in section 3.3. 2.

Following the analysis of burglary we consider in section 3.3.3 the analysis of a broader class of crime, *crime against property with violence.* As explained earlier we do this to model separately the trends in rural and urban crime. We also present a set of estimates of the determinants of the average value (in monetary terms) of property crimes in section 3.3.4.

3.1.1 Choice of variables

The choice of variables used in the analysis is clearly crucial. It is also particularly problematic in an exercise such as this since existing theory, economic and otherwise, gives only fairly weak guidance as to what should and should not be included. By contrast, for example, in looking at say the demand for money, there is a large body of empirical and theoretical research which places strong restrictions on the form of the model. Our strategy here has been guided by several key principles:

- *Practice in existing research*

  It makes sense to take into account best practice in comparable studies. Moreover a variable which determines crime in one country is likely to have a similar effect in another. We have been strongly influenced by the recent studies for Great Britain conducted by the Home Office (Field (1999) , Dhiri *et al (1999)*).

- *A desire to be as flexible as possible*

  Given the complex nature of crime and the dearth of quantitative research on the subject in Ireland it was considered prudent to allow a variety of variables to potentially influence crime. In the jargon of the literature we "let the data speak for itself".

- *Parsimony*

  Given the amount of information available (approximately forty years of annual data) there is a limit to what one can say. There is also an imperative to generate results that are interpretable. Both of these factors impose some constraints on the modeller. For example while one can think of many variables which might potentially affect crime there is a danger to including "too many". With too many variables and given the sample size available one's estimates become much less precise. Secondly, one may be interested in how the business cycle affects crime. In principle it might be argued that consumer spending and GNP could have distinct effects. However given that they are highly correlated it will be very hard to distinguish between the two.

After a certain amount of experimentation we isolated four variables as determinants of crime. One is demographic: the share of young males in the total population. This is consistent with a lot of international evidence and some for Ireland that young males are the group most likely to commit crime, particularly the sort considered here. There is one variable which reflects the general level of "prosperity" of the economy: real consumer spending per capita. Note that consumption was found to be the driving force in the recent UK study cited earlier. An alternative one could use which gives very similar results is GNP per capita inflation adjusted.

What is striking however is the absence of any labour market variable. Neither the level of wages nor the rate of unemployment appear in the model. This might seem unusual since the economics of crime emphasizes the importance of opportunity cost: as legitimate labour market activity becomes less rewarding the relative return to crime becomes more attractive. Nonetheless we were unable to detect a robust effect of the labour market on the level of crime *given the other variables included*. It is worth noting that a recent study for New Zealand (Papps and Winkelmann, 1998) cite evidence that just one half of studies of the crime/unemployment relation find a positive effect with the remainder finding no such effect.

There is an important issue that needs to be remembered: one should think of the macro-economy as generating at least two conflicting effects on crime. As an economy becomes prosperous, some criminals may switch away from illegal activities because the opportunity cost of engaging in crime is now higher. On the other hand there is also more to steal, so the rewards to a given crime are higher because peoples homes have more and better goods in them. Ideally one would be able to isolate these two effects but if they cannot be distinguished empirically what one observes is the "net effect" of the two opposing forces. An additional possibility is that as consumers have more money to spend they allocate additional resources to security in the form of burglar alarms and so on and this acts to depress burglary.

Finally we have two variables reflecting features of the criminal justice system: the average level of custody i.e. numbers in prison and detection rates of Group 2 crimes. Again it is important to note that controlling for one variable has implications for how we interpret the presence of absence of another variable. Therefore the fact that a measure such as the number of serving Garda Síochána does not appear in the model does *not* mean that they have no effect. Rather their effect works through the level of custody and the detection rates. Holding the latter constant it would be surprising if Garda numbers had an independent effect.

3.2 Econometric Issues

3.2.1 Modelling Strategy

We use a modelling strategy which is widely used in modelling aggregate time series data in a range of areas including crime trends. The framework needs to be flexible to allow for the intrinsically dynamic nature of the problem. In particular it needs to distinguish between short and long run determinants of the crime rate. It is possible that some of the variables we consider feature as long run determinants of the crime rate but not in the short run or vice versa. Details on the modelling strategy are contained in Appendix I to the report.

The model is built around the idea of an underlying equilibrium level of crime. We emphasize that equilibrium in this context has *no normative significance,* just as macroeconomic debates which refer to the "equilibrium level of unemployment" do not imply that there is a "desirable" level of unemployment. Rather there is believed to be a set of variables which interact to determine the level of crime. Permanent changes in the determinants act to shift the equilibrium level of crime. However there is no assumption that this adjustment takes place instantly. Indeed a common finding is that adjustment may take a considerable length of time. Temporary changes in one of the determinants of crime will cause a short run shift in crime away from its equilibrium level and we then expect the system to adjust back to the underlying equilibrium over time. Of course in the meantime one or more other variables may change causing additional perturbations to the system so the original equilibrium may never actually be re-attained.

The modelling strategy used here in closely tied up with an important body of econometric and statistical theory concerning estimation of time series relationships when some or all of the variables are *non-stationary*. In simple terms this refers to variables whose averages (and other characteristics such as variances) are not constant over time. Many macroeconomic variables such as GDP or wages are growing steadily over time unlike a variable such as interest rates which exhibit no long run trend. Considerable care is needed in the analysis of non-stationary series. There may appear to be a relationship between two (or more) variables that is, in fact, spurious. In simple terms two variables that are both following trends can appear to be related when there is in fact no causal relationship between them.

A consequence of this is that it is generally wise to analyse differences i.e. year-on-year changes and this, under specific circumstances, gets round the possibility of such a spurious regression. The procedure we follow is first to test for the existence of a set of variables between which there is a long run equilibrium relationship. If such a set of variables exist it is said that they are *cointegrated* - there is a linear combination of these variables that is possible. However identifying the long run equilibrium tells nothing about the dynamics of the relationship so one wants to also estimate a model that shows how changes in the determinants influence changes in the variable of interest. The whole process can be done in either a two step fashion or in one step. Since the results are very similar either way we present the two step method.

The modelling framework we use here is called an Error Correction Mechanism (ECM) and has been popularised by Hendry in a variety of macroeconomic applications. See Nickell (1985) for a clear exposition. The intuition for an ECM is that the dependent variable changes with changes in its determinants and where the system is relative to its long run or equilibrium level. If it has been temporarily "nudged" above its equilibrium level it will tend to adjust downwards over time, other things being equal and vice versa if it is below the equilibrium level so the deviation of the system from its long run level (the "error") is "corrected" over time. ECM's are now commonly used in modelling crime rates; in addition to the UK studies cited, recent work on Italy (Scorcu and Cellini 1998) and Germany (Entorff and Spengler 2000) use this approach.

3.3 Results

This part of the report will explain the basis for the model of theft and burglary, describe the model and give some results. Based on our discussion above the model is developed by estimating long run 'cointegrating ' relationships between recorded property crime and other factors, and (b) by estimating short run relationships which are corrected for the presence and effect of the underlying equilibrium level of crime

3.3.1 National Burglary Rates

3.3.1.1 Long-Run Relationship

Based on the technical issues discussed in Section 3.2 and Appendix I we first estimate whether or not the variables of interest can form a cointegrating regression. Table 3.1 outlines this first stage of the process, which can be thought of in its own right as estimating the effect of the independent variables on the long run equilibrium level of crime. This corresponds to equation (1) in the appendix.

Table 3.1: Estimating the long run relationship
Dependent Variable – Log of the National Burglary level

	Estimated Coefficient	*Standard Error*
Young Males as % Population (Log)	3.852*	0.542
Consumption (Log)	-0.297	0.260
Custody Rate (Log)	-0.886*	0.190
Detection Rate	-3.245*	0.303
Trend	0.054*	0.008
Constant	-78.865*	15.225
N	41	
Adjust R^2	0.993	

* indicates statistical significance of 90% or higher.

Interestingly the level of consumption does not have a statistically significant effect on the level of crime. What is evident however in this regression, as denoted by the asterisked terms, is that the long run or equilibrium level of crime is positively influenced by the stock of young males in the population - every 1% increase in the young male population (as a fraction of total population) raises burglary crime by 3.8%. Moreover crime appears to trend upwards, assuming all other things held constant, by about 5.4% per year. "Policy variables" such as the custody level and detection rate do tend to lower the level of crime. A 1% increase in detection rates or the numbers in custody lowers the level of burglary by 3% and 0.9% respectively.

We use the Johansen procedure to test for the existence of a long run relationship between these variables and are unable to reject the hypothesis that there is a unique cointegrating vector.

3.3.1.2 Dynamic Models

The long term relationship between crime, demography and the economy, as described above, determines what might be called the 'equilibrium' level of crime. There is nothing permanent or final about this equilibrium, which may change over time in response to demographic and economic changes. It may also respond to a range of other socio–economic factors or specific criminal justice variables.

Table 3.2 Dynamic Error Correction Models

		Standard Error
Growth in Young Male Population	4.720*	1.285
Growth in Consumption	-0.851*	0.323
Growth in Custody Number	-0.410*	0.127
Change in Detection Rates	-2.280*	0.471
Year	-0.000	0.0009
Error Correction Term (lag)	-0.650*	0.138
Constant	0.549	1.984
N	40	
Adjust R^2	0.7197	

Table 3.2 shows the results for the dynamic models which address these short-run influences but which also incorporate the error correction term from Table 3.1 to allow for the effect of the long run equilibrium relationship to be estimated. The Table shows a regression which relates the growth in burglaries to the growth in the independent variables. This corresponds to equation (8) in Appendix 1.

The key findings here are:

- Consumption growth tends to depress property crime growth – a 1% increase in the rate of growth in consumption lowers the growth in crime levels by 0.85%. This is consistent with a view that an improvement in the macroeconomic environment generates opportunities in the legitimate economy which raise the opportunity cost of crime. Note from Table 3.1 that this variable only has a statistically significant effect in the short run, there is no long run effect of macroeconomic prosperity on burglary.

- Changes in the growth rate of the young male population has a very large positive impact on crime – about 4.7% for every 1% increase in the rate of growth.

- The criminal justice system variables remain important negative determinants of the growth in crime.

- Finally the negative coefficient on the Error Correction term means that the system display sensible dynamic properties. That is when the level of crime is below its long run level it tends to rise until it attains equilibrium and similarly it falls when it is above the long run.

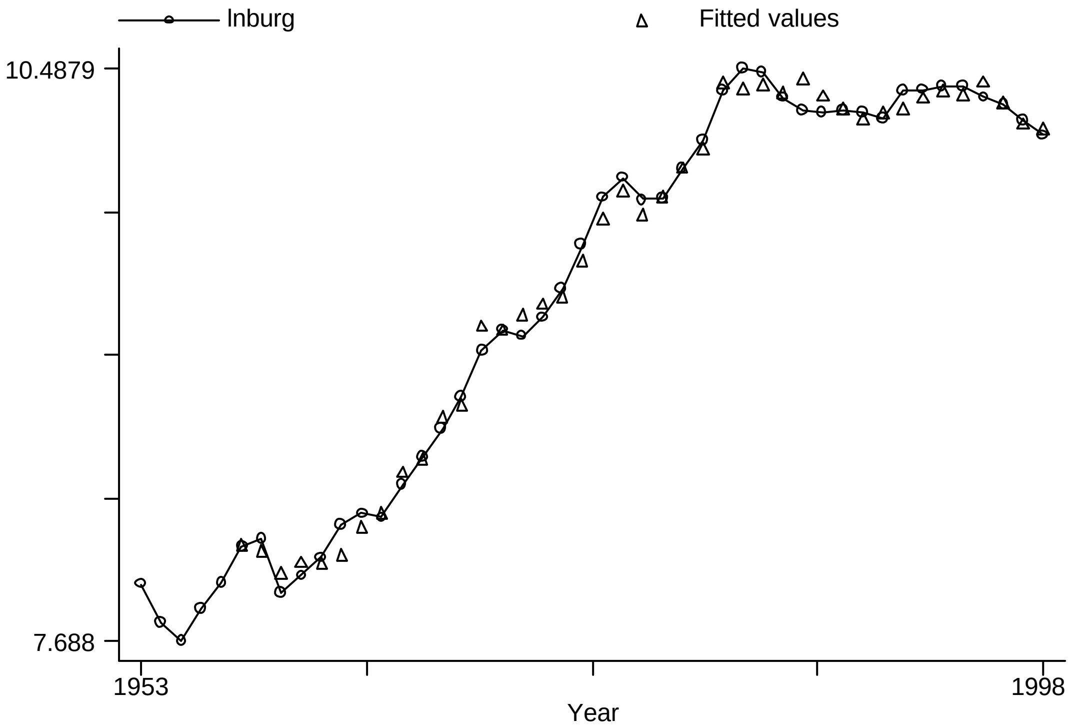
The results broadly accord with what theory and the international literature would suggest. For example Entorff and Spengler (2000) find the same effect of young males and the effect of deterrence for Germany as we find here. The importance of consumer spending is also found in a recent Italian study (Scorcu and Cellini 1998).

One issue, already raised in section 1, is the potential for criminal justice measures to be endogenous with respect to crime; that is criminal activity may influence policy responses. It is difficult to deal with this issue. As discussed in section 1 in order to do so we need to identify events which, for example, influence the detection rate but which themselves have no direct effect on crime[3]. These may only be available through some policy initiative which, for example, is piloted in a specific division or region in order to provide the research with a treatment group and a control group. One possible solution in the context of time series modeling is to use lagged values of the criminal justice variables in a procedure known as instrumental variables. This, when applied to the data used in table 3.1 and 3.2, shows no notable differences in the results. So this suggests, that insofar as we can control for this problem, that no major distortion is introduced by treating crime as exogenous.

3.3.2 Simulating National Burglary Rates

In this section we use the basic long-run relationship as illustrated in Table 3.1 as a basis for simulating a number of scenarios. In figure 3.1 we show the performance of the model in general in a plot of the actual data (labeled *lnburg*) and the fitted or predicted value for this based on the model in Table 3.1. We can clearly see that the model predicts the raw data very effectively – all the major turns in the raw series are matched in the fitted values and there are no major differences between actual and fitted data in any year.

Figure 3.1 Actual and Fitted Values – Long Run Relationship



The basic model can, for any values of the independent or explanatory variables, compute the corresponding value of the (log of the) burglary level. Therefore we need to be able to 'feed' values for the explanatory

[3] See Appendix 1 or Levitt (1996).

variables into the model. This allows some flexibility in the process from our perspective – we can choose values for the explanatory variables to simulate a particular underlying policy scenario for example.

Figure 3.2 plots the actual value for 1998 and the forecasts from 1999 to 2005 for a number of different scenarios.

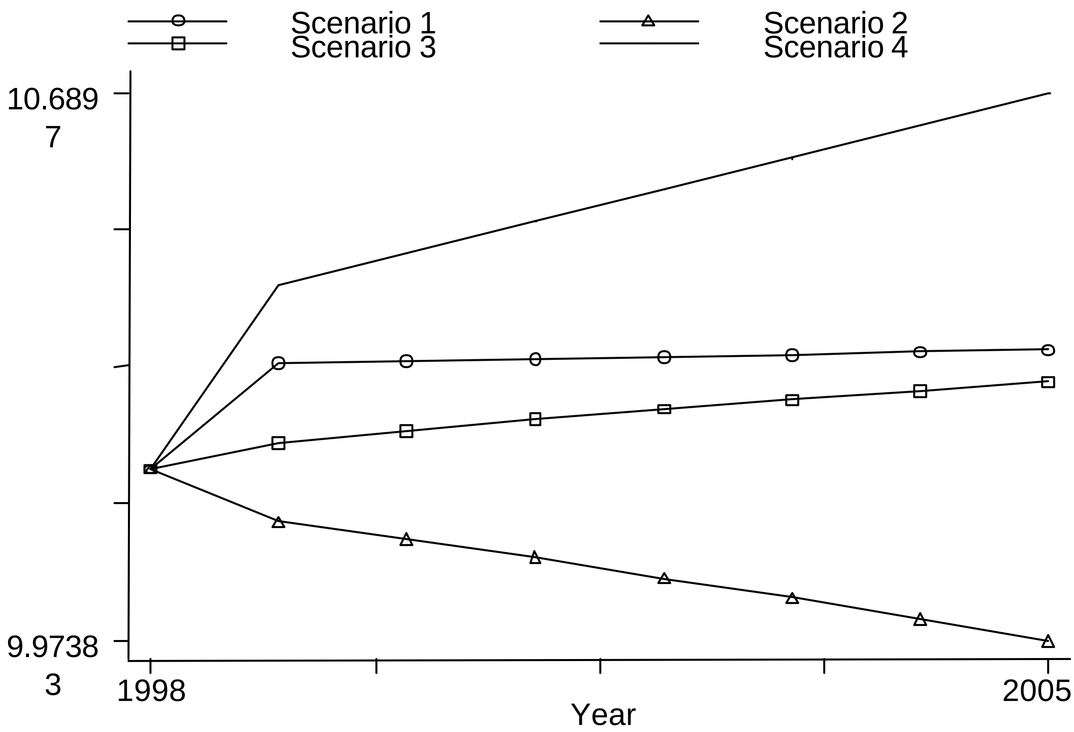
- In our initial scenario, Scenario 1, the values going forward in time of the explanatory variables are chosen on the basis of their individual past performance over the years 1980-1998. Simple linear trends are computed for each explanatory variable generating a forecast value for each independent variable. The underlying numbers here suggest a quite benign forecast for the share of young males in the population, a slight lowering of the detection rates and an increase in consumption and in custody rates. These are then used in the simulation. The outcome suggests a discrete blip upwards in burglary levels initially followed by a quite flat trend moving forward.

- Scenario 2 takes the same initial numbers as scenario 1 but replaces the previous assumption about detection rates with an assumption that detection rates *increase* by 2.5% per annum in each year. The strong negative effect of detection estimated in Table 3.1 is seen very clearly here. Even allowing for changes in all variables in line with recent trends there is a very strong downward shift in crime levels when this scenario is used in the simulation.

- Scenario 3 repeats the analysis of scenario 2 but also increases the share of young males in the population by 1% per annum. This compensates considerably for the negative detection rate effect and sees the simulated level of burglary returning to close to the outcome of scenario 1.

- Scenario 4 isolates the effect of young males on burglary by imposing the same 1% increase in the share of young men in the population per annum but allowing for all other variables to follow the trends as in scenario 1. This shows clearly the impact with a significant increase in the crime level effectively mirror-imaging the outcome in scenario 2.

Figure 3.2 Simulations 1999-2005: Four Scenarios



3.3.3 Regional Analysis

Table 3.3 shows the results of the regression model as applied to the rural and urban data separately. The dependent variable is the natural logarithm of G2 crimes owing to the unavailability of the burglary data by region as explained earlier, so comparisons with Table 3.1 and 3.2 should bear this in mind. Where relevant all explanatory variables are specific to the region under analysis.

The interesting issues in Table 3.3 are (a) the extent to which there are differences between the rural and urban results and (b) the differences between these results and the aggregate burglary figures in Table 3.1.

Comparing rural and urban we find that overall the results are very similar with the notable exception that the detection rate exerts a much more powerful effect in reducing crime in rural areas than urban. For every 1% change in the detection rate in rural areas the fall in crime is almost 3% compared to 1.3% in urban areas. Another interesting difference emerges when we look at the dynamic Error Correction model (in the bottom panel of Table 3.3). The effect of the growth in the young male population, so noticeable in the results in burglary, is shown to be only a feature of the urban sample.

Comparing these results with those for burglary in Table 3.2 we find that the qualitative results are very similar but there are important quantitative differences in the size of the effects. The negative effect of consumption on this class of crime, for example, is almost three times greater than on burglary.

Table 3.3 Rural/Urban Breakdown of Analysis

	RURAL		URBAN	
	Coefficient	Standard Error	Coefficient	Standard Error
Long run relationship				
Young Males as % Adult Population (Log)	0.3041	3.0097	1.5261	1.6837
Consumption (Log)	-0.9755	0.3997	-1.0996	0.2565
Custody Rate (Log)	-0.6985	0.3460	-0.4942	0.1472
Detection Rate	-2.7758	1.0680	-1.3144	0.2421
% Population in urban areas	-0.0320	0.1333	0.2768	0.0717
Trend	0.1038	0.0257	0.0148	0.0091
Constant	-186.5408	39.9605	-25.3805	12.2139
N	38		39	
Adjust R^2	0.9755		0.9956	
Error correction model				
Growth in Young Male Population	-3.3873	2.4175	4.4604	1.4800
Growth in Consumption	-1.2208	0.5563	-1.1069	0.3208
Growth in Custody Number	-0.1019	0.2097	-0.2776	0.1216
Change in Detection Rates	-1.0529	0.5250	-1.3274	0.2907
Change in % Population in urban areas	-0.0528	0.2399	0.1808	0.1843
Year	-0.0027	0.0031	0.0001	0.0027
Error Correction Term (lag)	-0.3639	0.1750	-0.6826	0.1554
Constant	5.4295	6.1844	-0.1701	5.4055
	38		38	
	0.4113		0.7806	

3.3.4 Estimates Based on the Value of Crime

An alternative way of modeling crime is to consider the value of a crime defined as the total value of goods taken divided by the corresponding number of crimes. Table 3.4 shows the results from our two-step model with the dependent variable now based on this average inflation adjusted value of crime data (as opposed to the number of crimes).

Table 3.4 Estimates based on Average Value of Crime

Long run relationship	Coefficient	Standard Error
Young Males as % Adult Population (Log)	2.0865*	1.2440
Consumption (Log)	0.9284*	0.4336
Custody Rate (Log)	-0.7717*	0.2813
Detection Rate	-1.7818*	0.5533
Trend	0.0117	0.0154
Constant	-15.1002	29.5715
N	41	
Adjust R^2	0.9399	
Error correction model	Coefficient	Standard Error
Growth in Young Male Population	0.5941	2.6880
Growth in Consumption	1.6041	0.4277
Growth in Custody Number	-0.5436	0.3231
Change in Detection Rates	-2.2862	0.5951
Year	-0.0002	0.0021
Error Correction Term (lag)	-0.6211	0.1466
Constant	0.4430	4.2493
N	40	
Adjust R^2	0.487	

In the long run relationship we see results that are broadly in line with the earlier findings based on the number of crimes. The most notable difference is that consumption is a significant and positive determinant in both the short run and long run. This contrasts with the earlier finding that consumption only has a negative and short run effect when looking at the number of crimes. This implies that pro-cyclical swings in consumer spending generate fewer but more lucrative crimes. This suggests that it is predominantly “small time” crime which is reduced when the macroeconomy improves. As the economy improves those involved in high value crime benefit from the greater prosperity in that individual crimes are more lucrative. On the other hand those involved in low value crime may be attracted away to the legitimate economy hence the finding of a positive effect on the value but not on the level. Also growth in the young male population does not have a significant effect on the value of crime.

3.3.5 Alternative specifications

A number of alternative models were estimated throughout the completion of the report, which, for reasons of brevity, have been omitted. However some important points are worth noting about these specifications. One variable that some will be surprised to see omitted from the list of explanatory variables is unemployment or some other direct measure of labour market activity such as the wage level. However, once we control for the explanatory variables used in the modelling, we were unable to find *any* robust effect from these labour market variables. Given the mixed evidence on this in the international literature it is perhaps not that surprising. We also experimented with migration rates and with the level of burglary in the UK however in both cases we could eliminate them from the model.

4. Conclusions

In this report we present an econometric analysis of the trend in Irish burglary crime between 1958 and 1998. We have collected an extensive dataset based on Garda reports and other official publications and we use this data to develop an econometric model of burglary and the wider category of crimes against property with violence (known as Group 2 crimes in the published statistics). The model explains burglary in terms of demographics, the macroeconomy and two characteristics of the criminal justice system. We distinguish between the determinants of the level of crime (the long run or equilibrium level) and the short run dynamic behaviour of crime. We develop the model for national data and also for an urban/rural split. Moreover we consider data on the value of property stolen in Group 2 crimes as an additional indicator of crime level.

Key findings are:

- Macroeconomic conditions as measured by the level of real consumer spending per capita have little effect on the equilibrium level of national burglary crime but, in the short run, growth in consumer spending does depress crime. However consumption has a much more pronounced and positive effect on the average value of a burglary act. Hence the evidence suggest that consumption affects the nature of the crime but not the overall frequency of crime in the long run.

- Demographic conditions are measured by the share of young males in the adult population. This variable has a positive effect on crime both in the long run and the short run.

- The criminal justice system impacts on crime through two channels – the detection rate (reflecting the probability of being caught), and the numbers in custody for Group 2 crime (reflecting the consequences of being apprehended). Both have well determined negative effects on crime reflecting the deterrence effect and the direct effect of incarceration.

- Some difference can be noted when we consider the data separately for rural and urban populations. Detection has twice as big an effect on crime in the rural areas possibly reflecting the more risk averse nature of criminals in these areas.

- One somewhat surprising result is that we were unable to find any robust effect from direct measures of labour market activity such as unemployment rates or wage levels.

[redacted]

5. References

Bacon Peter and Martin O Donoghue (1975) "The Economics of Crime in the Republic of Ireland: An Exploratory Study" *Economic and Social Review*, 7,19-34

Becker, Gary (1968). "Crime and Punishment: An Economic Approach." Journal of Political Economy, 76(2), March-April, 169-217.

Central Bank of Ireland, (1977). Technical paper no. 1/RT/77.

Corman, Hope and Mocan, H. Naci (2000). "A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City." American-Economic-Review; 90(3), June, pages 584-604.

Dhiri Sanjay, Sam Brand, Richard Harries and Richard Price (1999) "Modelling and predicting property Crimes trends" Home Office, London (Research, Development and Statistics Directorate, Research Study 198

Entorff Horst and Hannes Spengler (2000) "Socioeconomic and demographic factors of crime in Germany" International Review of Law and Economics, 20, 75-106.

Field Simon (1999) "Trends in Crime Revisited" Home Office, London (Research, Development and Statistics Directorate, Research Study 195

Fitzgerald John , Jonathan Hore (2001) ESRI Times series databank ESRI Dublin

Kelly, Morgan (2000). "Inequality and Crime." Review of Economics and Statistics, 82(4), 530-39.

Leddin Anthony, Brendan Walsh(1998) , Macroeconomy of Ireland (4th ed.). Dublin: Gill and Macmillan.

Leddin Anthony, personal communication, (January 2001)

Levitt, Steven D. (1995). "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" National Bureau of Economic Research Working Paper: 5268, September.

---- (1996). "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation." Quarterly-Journal-of-Economics; 111(2), 319-51.

---- (1997). "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." American-Economic-Review; 87(3), June, pages 270-90..

McCullagh Ciaran (1992) "Unemployment and imprisonment: examining and interpreting the relationship in the Republic of Ireland" *Irish Journal of Sociology* , 2, 1-19

Nickell Stephen (1985) "Error Correction, Partial Adjustment and All That: An Expository Note" Oxford Bulletin of Economics and Statistics; 47(2), 119-29.

Papps Kerry & Rainer Winkelman (1998) "Unemployment and Crime: New Answers to an old question", IZA (Bonn) Discussion paper no.25

Pudney Stephen, Derek Deadman and David Pyle (2000) "The relationship between crime and punishment and economic conditions: is inference reliable when crimes are under-recorded?" *Statistics in Society* (Journal of the Royal Statistical Society, series A) , 163, 81-98.

Rottman David B. (1980) *Crime in the Republic of Ireland: statistical trends and their interpretations* ESRI, Dublin ; General Research Series no. 102.

Rubenstein, Edwin and Robert Woodson (1995). "Crime in America: Fighting Back with Moral and Market Virtues." IMPRIMIS, August 1995, Vol. 24, No. 8.

Scorcu Anonello and Roberto Cellini (1998) "Economic Activity and crime in the long run: an empirical investigation on aggregate data from Italy 1951-1994" International Review of Law Economics, 18, 279-292

Watson Dorothy (2000) Victims of Recorded Crime in Ireland: Results from the 1996 Survey ESRI/Oak Tree Press

"Whittaker Report" (1985): *Commission of Inquiry into the Penal System Report* (chairman T.K. Whittaker), Dublin Stationery Office

Winter-Ebmer Rudolf , Stephen Raphael (1999) "Identifying the Effect of Unemployment on Crime'" CEPR (London) Discussion Paper no. 2129

Appendix I - Introduction to multiple regression

This is a simple introduction to econometric models and serves to introduce some of the concepts and terminology used in the report. An econometric model seeks to explain one or more variables in terms of a set of variables. For simplicity we consider *single equation models* where there is only one variable to explain. The variable to be explained is called the dependent variable and is frequently denoted by "Y". The variables explaining it are variously described as the independent or explanatory variables and usually denoted with "X's.

$$Y = b_o + b_1 X_1 + b_2 X_2 + h \qquad (1)$$

The parameters to be estimated are the β's , the first of which is the constant or intercept. β_1 is the marginal effect of a unit (1) change in X_1 on Y *holding any other variables in the model constant* (X_2 in this case). This conditionality is important to remember since in social sciences one is typically analyzing a variable which has several determinants.

The final term, η, is a random error term and reflects the fact the data points will seldom, if ever, lie exactly on the estimated line. This model is described as "linear" because the parameters to be estimated appear linearly in the equation. This does not constrain the relationship between Y and any of the X variables to be linear since one can include non-linear functions (or transformations) of the X's as independent variables. Say one believed on theoretical grounds that the relationship between Y and X_2 is "increasing concave", that is it has a positive but diminishing effect. Including the square of X_2 and hence an additional parameter β_{22} allows this to happen. We would expect β_{22} to be negative and β_2 to be positive.

$$Y = b_o + b_1 X_1 + b_2 X_2 + b_{22} X_2^2 + h \qquad (2)$$

The actual values of the parameters that are estimated from the data are usually represented by $\hat{\beta}_1, \hat{\beta}_2$ etc. Strictly speaking (1) or (2) is a statement about the true relationship between a group of variables in the *population* and we use data from a *sample* to infer estimates of those parameters. A convenient property of regression models is that they can be interpreted as explaining deviations in Y from its mean as a function of deviations of the explanatory variables from their means (a "bar" over a variable denoting the respective mean). For the model in (1) above this implies:

$$Y - \overline{Y} = \hat{b}_1.(X_1 - \overline{X}_1) + \hat{b}_2.(X_2 - \overline{X}_2) \qquad (3)$$

It is quite common for the dependent variable to be a transformation of the underlying variable of interest. Specifically, rather than using Y itself, one replaces it with its natural logarithm (usually denoted LogY or lnY). There are two principal reasons for this, statistically this specification tends to give a better "fit", that is it is easier to explain the logarithm of a variable rather than its level. More importantly, the parameters now give the *proportionate* impact on Y so in the model below (4), if β_1 was estimated to be 0.5 then a unit change (an increase of 1) in X_1 would cause Y to rise by approximately 50%.

$$\ln Y = b_o + b_1 X_1 + b_2 X_2 + h \qquad (4)$$

In many cases researchers will utilise a "double-log" specification where one or more of the independent variables is also represented by its logarithm:

$$\ln Y = b_o + b_1 \ln X_1 + b_2 X_2 + h \qquad (5)$$

The advantage of this is that the estimate of β_1 corresponds to an *elasticity* : a measure of the proportionate response of the dependent variable for a 1% change in the independent variable. Say X_1 above is the level of GNP in the above model. Then if we estimate β_1 to be 0.4 this implies that a 1% increase in GNP (as distinct from a 1 unit increase) causes a 0.4% increase in the dependent variable.

In the literature on crime the vast majority of studies use this transformation of the level of crime and we follow this approach. In some cases researchers have defined the level of crime as the ratio of the number of crimes to the level of population. This is, to some extent, a matter of taste and is unlikely to be of importance.

In econometric models, the independent variables should be "exogenous" that is they are determined outside of the model and in particular should not be themselves influenced by the dependent variable in which case the X variables will be correlated with the error term η. If this does not happen then this causes the estimated parameters to be biased. This is what is known as *endogeneity* or *simultaneity bias* .

In essence what the social scientist is trying to do is to mimic what the natural scientist does in analysing experiments. He/she administers a treatment (say a drug) to a randomly chosen group of patients and compares them with a comparison group who on average are the same in every other respect. If the treatment group differed systematically from the comparison group either because the experimenter chose to give the drug (say) to the more seriously ill or individuals selected themselves into the groups on some basis then it is not hard to see that a simple comparison will probably give a misleading estimate of the drug's effect. This is why a “double blind” experiment where neither the subject nor the scientist knows who is being treated is preferred practice.

A common solution to this problem in the econometrics literature is the use of Instrumental Variables. This involves finding some other variable, say Z, which does not directly affect Y but is correlated with the X variable that is endogenous. One can use the relationship between the instrument, Z, and the offending X variable to "purge" it of its endogeneity. Identifying an appropriate Z variable to use is partly a matter of luck and partly judgement. There is no guarantee that one will exist. This procedure also, under certain conditions, can be used to control for the *attenuation bias*, caused by measurement error.

Careful use of the appropriate statistical tests allows one to identify the problem and test the adequacy of the solution. In some cases one can be confident that a particular variable will not be endogenous. In a model of crime where the level of GDP is an explanatory variable, it seems highly implausible that there would be a noticeable feedback from the level of crime to the level of GDP though it could effect the output of certain sectors (like the security industry). When the independent variable is under the control or influence of the government things are less clear. In the case of crime, say we are interested in the impact of the level of policing on crime, it is reasonable to expect a negative relationship. The problem is that the level of policing may be determined by government as a response to increasing crimes. This is the reverse causality often discussed. This generates a positive relationship between the two. Estimation of the model by standard methods (Ordinary Least Squares) produces an estimate of a parameter which combines two effects but gives us an estimate of neither. If we can find another variable which directly effect the level of policing but not crime directly then this serves as an Instrumental Variable which allows to estimate the parameter of interest in a statistically satisfactory way.

The data used in study consists of annual observations on different variables. This is called *time series* data as distinct from *cross section* data where one looks at a survey of individuals or families or firms. In dynamic models, time enters explicitly in the model. Analysis of economic time series requires particular techniques of analysis and these have developed rapidly in recent years. Consider the following model:

$$Y_t = b_o + b_1 X_{1,t} + b_{11} X_{1,t-1} + b_2 Y_{t-1} + h \qquad (6)$$

here Y in any given year t depends on the current and last years value of a variable X_1 as well as the variable's own previous value Y_{t-1} . The coefficient on the latter, (called the lagged dependent variable) β_2 , determines how slowly or otherwise Y responds to a change in X. It would typically be between 0 and 1, with values close to 1 generating very gradual adjustment. In (6) the presence of the previous years value of X means that a given change in the variable generates an "aftershock" : it has both an immediate effect and a lagged effect.

The modelling framework we use here is called an Error Correction Mechanism (ECM) and has been popularized by David Hendry in a variety of macroeconomic applications; see Nickell (1985) for a clear exposition. This has generally been found to provide a good fit for many series and is quite intuitive. It turns out to be closely related to the question of *cointegration* due a celebrated theorem of Granger and Engle that shows that an ECM can represent any set of cointegrated variables. For simplicity consider where the equilibrium level of Y depends on two variables X and Z.

$$\boldsymbol{Y_t} = \boldsymbol{a}_0 + \boldsymbol{a}_1 \boldsymbol{X_t} + \boldsymbol{a}_2 \boldsymbol{Z_t} \qquad (7)$$

An Error correction mechanism takes the form[4]:

$$\Delta \boldsymbol{Y_t} = \boldsymbol{b}_0 + \boldsymbol{b}_1 \Delta \boldsymbol{X_t} + \boldsymbol{b}_2 \Delta \boldsymbol{Z_t} - \boldsymbol{b}_3 (\boldsymbol{Y}_{\boldsymbol{t}-1} - \boldsymbol{a}_0 - \boldsymbol{a}_1 \boldsymbol{X}_{\boldsymbol{t}-1} - \boldsymbol{a}_2 \boldsymbol{Z}_{\boldsymbol{t}-1}) \qquad (8)$$

$$\Delta \boldsymbol{Y_t} \equiv \boldsymbol{Y_t} - \boldsymbol{Y}_{\boldsymbol{t}-1}$$

The term in parentheses, the Error Correction term, represents deviations of y from its long run. Hence changes in Y can be attributed to changes in the underlying determinants (ΔX, ΔZ) or the system being away from its equilibrium level. The term in front of the parenthesis is negative (assuming , as is normal, that $b_3>0$) representing the assumption that if the system is away from its equilibrium level at any point it will tend to revert back over time the "error" is "corrected". If however one has identified y, x and z as being co-integrated then one obtain estimates of the parameters of the long run relation (a_0, a_1 and a_2) by first estimating (7) and then substituting the lagged residuals into a models in first differences. Estimating that gives one estimates of the parameters of (2): That is one estimates:

$$\Delta y_t = b_0 + b_1 \Delta x_t + b_2 \Delta z_t - b_3 (y_{t-1} - \hat{a}_o - \hat{a}_1 x_{t-1} - \hat{a}_2 z_{t-1}) \qquad (9)$$

This is the two step procedure that we shall use. An important feature of this class of model is that there is a distinction between the long and short run effect. The system adjusts gradually to a change in one of the determinants.

Another, related, reason for estimating a model in the form of (8) or (9) is to do with a feature of time series data, *stationarity*. Broadly speaking a series is stationary if its key characteristics such as the mean and variance do not change over time. That is the series fluctuates over time round a constant mean. Such a series is denoted I(0). However many economic time series are not stationary since they are trending, usually upwards, over time. For example GNP, labour productivity or real wages are all rising over time though they may fall in particular periods. Hence their means are changing over time. Estimation of models with such non-stationary series is not straightforward and great care needs to be taken since there may appear to be significant relationships between series when there is in fact none - this is called a *spurious regression*.

[4] There are actually several ways of writing an ECM for this system, this is the most intuitive for the present purposes.

It may be the case that the year-on -year change in the series is stationary. For our dependent variable this means that $\Delta Y_t \equiv Y_t - Y_{t-1}$ is I(0), in which case Y_t is denoted I(1). That is the series has to be "differenced" once to achieve stationarity . If it has to be differenced twice to achieve stationarity then it is I(2) and so on. If a linear combination of I(1) variables are I(0) then they are said to be cointegrated. This means that there is a long run relationship between them (which may not be unique).

There is some discussion in the international literature over whether theft and burglary are I(0) or I(1). The balance of evidence in the literature suggests that crime variables are I(1), or stationary only for the year-on-year difference. Formal testing of this for the data used here (using what is known as the Augmented Dickey Fuller test) confirms this[5].

Finally, a brief word on under-reporting is appropiate. In econometrics and statistics this is classified as "measurement error" and there is an extensive literature on the subject. One can briefly state a number of key results. If the measurement error in a variable is non-random then this is likely to give rise to misleading estimates of parameters, the estimate is likely to be above or below the true value. For random measurement error (or "classical measurement error" as it is known) one can state two results. If it arises in the dependent variable (the "Y") then there is *no bias* to ones estimated parameters 'though one's estimates will be less precise. This may seem surprising since the dependent variable is in a sense the most important one, however it is straightforward to show this. Random measurement error in an independent variable (an "X") causes a *downward* bias in one's estimates of the effect of this variable - so called "attenuation bias".

[redacted]

5 The non parametric Phillips-Perron test gives the same results in general.

Appendix II - Data Appendix

The following table describes that variables used in the econometric analysis. As well as a description of each of the variables, we also list the source and any other relevant information. We do provide a detailed description of data series which we considered at an early stage in the analysis but which do not feature in the analysis here. A brief description is given at the end of this section.

Burglary

Description:	The total annual number of indictable offences for burglary.
Source:	Annual Report of the Commissioner of the Garda Síochána. Dublin, Stationary office (various years).
Notes:	(i) The twelve-month period covered by the annual report is not consistent over the fifty-two years of the data. Specifically, the series is recorded as follows: 1947-1957, December-to-December calendar year; 1958-1974, September-to-September; 1975-1998, December-to-December; and 1999, January-September. The change in 1999 is entirely due to the introduction of the national police computer, 'PULSE', in that year. Since the data for 1999 would clearly not be comparable we only use data up to and including 1998.
	(ii) The burglary series is a subset of the Group II crimes – i.e., all crimes against property where violence was used. However, what exactly should be counted as burglary is not clear. Up until 1976 the group II category of crime covers several sub-groups, these include: burglary, sacrilege, house-breaking, breaking (and attempts to break) into shops, warehouses and other premises, entering with intent to commit a felony and possession of house-breaking tools. From 1977 onwards, however, all of these categories, with the exception of burglary, are not reported individually. After careful analysis of the reports, we concluded that the category labelled burglary from 1977 onwards actually includes all of the categories that were previously reported individually. Therefore, in constructing the series that we call 'burglary' for the pre-1977 period we summed all of the above categories that were reported individually. This, we believe, makes our pre-1977 definition of burglary consistent with the definition of burglary that is listed in the Garda reports after 1976.

Group II

Description:	The total annual number of indictable offences for 'Group II' crimes (violent crimes against property) between 1947 and 1999.
Source:	Annual Report of the Commissioner of the Garda Síochána. Dublin, Stationary office (various years).
Notes:	As is the case with all data obtained from the Commissioners'

	Reports, the twelve-month period that the series covers is not consistent (see note (i) for Burglary).
	Group II crimes can be further broken down into those crimes committed in the five districts Dublin (DMA), Cork, Limerick, Galway, and Waterford. Throughout the years, there are other regional breakdowns of the series, though none are consistent. We aggregate these five divisions to form an "urban" series and calculate a "rural" series by subtracting this series from the aggregate.

Total value of property stolen in violent crimes against property

Description:	This is the total value of all property stolen in burglaries and robberies where violence was used. The group II categories covered by this value series are sacrilege (including larceny from donation boxes), burglary (according to the original definition in the Garda reports – see note (ii) above for the burglary series), housebreaking (dwelling houses) and breaking into shops warehouses.
Source:	Annual Report of the Commissioner of the Garda Síochána. Dublin, Stationary office (various years).
Notes:	(i) See point (i) for Burglary above.
	(ii) The number of cases used to calculate the value series is not equal to the total number of group II crimes for two reasons. Firstly, for reasons that are not clear to us, it does not include all of the categories that are in group II (see the description above). Secondly, it excludes a number of indictable offences where nothing was stolen. The annual Garda reports do, however, list the number of cases corresponding to the value series, and from this we can work out the average value of property stolen per crime committed.
	(iii) There is no value series for the year 1972, we therefore use the existing data to calculate a value consistent with the trend. There are several outliers in the value data which we considered it necessary to adjust. These occur in 1974 , 1983 and 1986. For 1974 and 1986 the Garda report explains that the series in those years is affected by particular thefts of art work to the value of £5 million and £9.289 million respectively. We adjusted the data by simply subtracting these values. We were unable to discover the reason for the "spike" in the data in 1983 - the data for 1982 to 1984 are : £18.5m, £31.8m 21.7m respectively. We simply replaced the 1983 value with the mean of 1982 and 1984 values. More complicated methods of imputation are available to deal with this problem but we are confident that it would make no difference.

Total value of property recovered from violent crimes against property

Description:	Total value of the property recovered from the preceding series
Source:	Annual Report of the Commissioner of the Garda Síochána. Dublin, Stationary office (various years).
Notes:	See notes (i) and (ii) for the series total value of property stolen. The data for 1974 was adjusted in the light of the "spike" in the data. A explained above a single crime to the value of £5million generates an outlier for the 1974 value of property stolen . This is also reflected in the "Value recovered" series i.e. there is a blip in that series also so we simply subtract the £5m from this series as well.

Group II detections/detection rates

Description:	The number of indictable offences in group II detected by the Gardaí.
Source:	Annual Report of the Commissioner of the Garda Síochána. Dublin, Stationary office (various years).
Notes:	See notes (i) and (ii) for Group II above. In addition to the aggregate data on detection we calculate series under two geographical headings "rural" and "urban" . The later consists of five Garda districts Dublin, Cork, Limerick, Galway and Waterford. The rural figure is calculated by subtracting the urban data from the aggregate figure.

Daily average numbers in custody

Description:	The daily average number of prisoners in custody in all institutions (including juvenile) in the state.
Source:	Statistical Abstract, Central Statistics office (1947-1997), which covers the period up to and including 1994. Data for 1995-1998 is taken from the Irish Prison Services Report on Prisons and Places of Detention for the Years 1995-1998, Table1.
Notes:	A breakdown on the basis of gender is available up to 1979, after which point it is not reported.

Demographic variables

Description:	These are the variables we use to calculate the number of crimes per head of the population, such as in figures 2A and 3A. The following age breakdowns are available for both males and females: aged 10-14, 15-24, 25-64 and 65+. For "Young males" we take the series corresponding to Males aged 15-24. In the analysis this is divided by total population , the sum of all age categories for males and females.
Source:	The data is taken from Table 30 of the ESRI Databank of economic time series for Ireland .
Notes:	This is the same data that the ESRI use in constructing its Medium Term Model for the Economy. The demographic variables are available from 1953 onwards.

Consumption per capita

Description:	This is the level of personal consumer spending in constant prices (£ millions) divided by total population (000's).
Source:	The consumption data is taken from Table 6 of the ESRI Databank of economic time series for Ireland. The population figure is as described above.
Notes:	This is the same data that the ESRI use in constructing its Medium Term Model for the Economy.

Other variables:

Data on Garda numbers are taken form the Statistical Abstract of the Central Statistical Office.

Data on British burglary levels were provided by the Home Office and were those used in Home Office Research Study 198 (Dhiri *et al* 1999)

Migration
Data on migration were taken form the ESRI databank, Table 30.