Section B

Non-linear modelling of burglary and violent crime in the UK

Paul Ormerod, Craig Mounfield and Laurence Smith
Volterra Consulting

Contents

Summary of non-linear modelling of crime study

The aim of this project is to offer an innovative framework in which to understand better the process by which crime rates spread (or contract) over time. The approach is similar to that used in mathematical biology to describe how potential epidemics are either spread or contained in a population. A key motivation for this is the fact that epidemics are spread by social contact; by social interaction between individuals. In the same way, participation in crime can be influenced by the activities of individuals' peer groups.

It is a well established fact in research literature that the statistical distribution of criminality among the population is highly skewed: many people commit no crimes, the typical criminal commits very few crimes, and there is a long statistical tail stretching out to a small minority who commit many. Consequently it is justifiable to model a population at any point in time as if it were divided into a small number of discrete groups which differ in their potential to commit crime. There is a set of flows between these groups, whose overall effect describes the evolution of crime rates. The flows are postulated to depend upon factors such as the severity and efficiency of the criminal justice system and general social and economic conditions.

The study examines the recorded levels of burglary and of violence against the person. There is an important debate about the reliability of this data, particularly that for violent crime. It is not in the terms of reference of the project to enter into this debate, although an initial amount of work shows that the model can be applied successfully to the British Crime Survey data on violent crime, which includes unreported crime), as well as to the recorded violent crime data.

We draw on a wide range of evidence from the literature on crime to calibrate our models against the levels of crime that were obtained in the late 1990s. In other words, solutions of the model are obtained which give these levels of crime.

The model is used to analyse what the effect of changes in policy on recent levels of crime might be. We examine the effect of:

- doubling the average prison sentence
- doubling the probability per offence committed of going to prison
- reducing dramatically the rate of recidivism.

We examine these changes both singly and in combination, with and without changes in the overall level of deterrence to accompany them.

An important implication of the model is that, whilst such changes in the criminal justice system do have impacts on crime, these are not large, because the process of creating criminals is not directly affected by these measures. By contrast socio-economic forces, such as the quantity of goods available to be stolen, or relative economic deprivation (which enhances the incentive to steal them) can and do create criminals and may therefore have substantial effects. The empirical crime literature gives no firm indication of the impact of deterrence, although on balance it might be thought to be more likely to arise from changes in the probability of being caught rather than by sentencing policy. The model shows that if deterrence changes along with changes in the punitive aspect of the criminal justice system, the impact on crime is much greater than it is with changes in the latter factors alone.

Social and economic conditions are important, and the model suggests, for example, that the rise in inequality over the past 20 years or so has had an important impact on crime. This is not to say, of course, that individuals are motivated directly by inequality. But rather, as the survey by Robin Marris points out, it is the overall circumstances associated with poverty which can lead to an increase in crime.

Social interactions between individuals – the so-called 'fashion' element of crime – play a strong role, and changes in the 'fashionability' of crime amongst the main perpetrators, young relatively unskilled men, can have a substantial impact. This sort of factor is difficult to measure in practice, and its effect may well be a reason why formal forecasting models of crime such as those built within the Home Office often have a poor track record of forecasting in practice.

1. Introduction

In the past half century crime rates increased dramatically in many Western countries and have begun to fall back only in the past decade. Consequently crime has become a major issue of public policy especially in the Anglo-Saxon countries. A very large literature has developed on the general causes of, and the impact of public policy on, crime. Yet no consensus has emerged on many quite basic issues.

Economic analysis of the phenomenon of crime was stimulated in the late 1960s by the paper by Becker (1968). In this framework, agents in the market for crime – for example, criminals and the law enforcement agencies – are assumed to act in accordance with the rules of optimising behaviour. They are able to both gather and process substantial amounts of information efficiently in order to form expectations on the likely costs and benefits associated with different courses of action, and to respond to incentives and disincentives in an appropriately rational manner

But amongst research carried out within the framework of economic theory, few clear conclusions emerge. Ehrlich (1996), in an excellent exposition of the approach to crime based upon such theory notes that the empirical literature is 'voluminous'. However, he states that on the crucial question of the impact of positive and negative incentives on crime, 'it would be premature to view the empirical evidence as conclusive'. Ehrlich notes that the quantitative estimates of such effects vary, even to the extent of a minority of studies failing to find any effect at all, and part of his paper is devoted to a discussion of the potential reasons for this, not least of which is the intrinsic limitations of crime statistics. Indeed, an awareness of the often seriously unreliable nature of the statistics is one thing on which most economists and criminologists find common ground.

The survey by Marris (2000) draws on both the criminological and the economic literature, and is probably the most comprehensive which has been undertaken, citing almost 300 references. Marris reinforces the above points strongly. As he points out, a wide range of causes of crime has been suggested in the theoretical literature, particularly for property crime. These vary from the macro-economic environment and factors such as real average incomes, unemployment and inequality, to characteristics of individuals given by factors such as their family structure and level of intelligence. A substantial group of potential influences relates to the criminal justice system and factors such as sentencing policy, the probability of catching a criminal, and possible deterrence effects.

However, Marris concludes that *'it cannot be stressed too strongly that the existing literature, voluminous as it may be, does not provide any firm, unequivocal guidelines based upon empirical evidence on the true causes of crime'*. He says that neither criminology nor economics can fully explain the 30 year Anglo-Saxon crime wave that began in the mid 1950s, nor can these disciplines explain the subsequent downturn. Marris puts forward two reasons. First, only a modest fraction of the list of theoretical influences has been tested empirically in the literature. Second, results on the impact of the factors which have been tested vary very substantially.

The reasons why empirical results obtained differ so much are complex. Sometimes, they are particular to individual studies, where the methodology used has not necessarily been in line with modern best practices. However, even within the context of careful research carried out on modern lines, the results still vary extensively between studies. Even within a single study, the results which are obtained are often very sensitive to the particular set of data which is used. The addition (or subtraction) of a single extra observation from the data set used can make distinct quantitative changes to the results. This point was made clear in Ormerod and Smith (2000).

The aim of this paper is to offer a different methodological approach to the process by which crime rates spread (or contract) over time. The approach is similar to that used in mathematical biology to describe how potential epidemics are either spread or contained in a population (see, for example Murray,1990). We divide the population at any point in time into a small number of discrete groups which differ in their potential to commit crime. There is a set of flows between these groups, whose overall effect describes the evolution of crime rates. The flows are postulated to depend upon key elements identified in the literature, such as the impact of incentives or general social and economic conditions.

An essential element in this model is social interaction between individual agents. The model is developed at the aggregate level, describing the outcome of decisions made by individual agents of whatever degree of (bounded) rationality might be thought appropriate. Individuals form views on what we might call external factors, such as the overall social and economic conditions and the punishment structure, and use these to determine their movement or otherwise in or out of the different categories of agent in the model. But they are also influenced by the behaviour of others.

Chapter 2 sets out in some detail the development of the formal model. Chapter 3 discusses the application of the model to post-war data on recorded burglaries in the UK, and Chapter 4 applies the model to the recorded data on violence against the person.[1] Chapter 5 sets out some policy implications, and Chapter 6 presents the conclusions.

[1] The use of recorded crime data was a clearly defined aspect of the work commissioned by the Home Office. There are obviously, however, important issues concerning the meaning of such data, especially with regards to violent crime. The interpretation of the recorded crime data is, however, outside the scope of this project.

2. The formal model

Overview

Initially, we can think of the population as being divided conceptually into three groups. First, those who are not susceptible to commit a crime (denoting this group subsequently as N, for not susceptible). In other words, individuals with a zero probability of committing a crime. As a not unreasonable approximation, for example, most women, certainly those over 25, might be placed in this category, as might most pensioners. We should stress, however, that the word 'population' does not necessarily refer to the population of an entire nation, and could equally well apply to that of an individual city or even a neighbourhood within a city.

The second group is made up of the susceptibles (S), those who have not yet committed a crime, but might well do so, or those who commit crime on an occasional basis. It is very well documented in the criminology literature that young men in their teens and early 20s are particularly prone to commit crimes (see, for example, Austin and Cohen (1995), DeIulio (1996), Marris (2000)). Of course, by no means all men in this age group actually commit crimes, but there is a high propensity to do so, from acts of minor vandalism carried out in what used to be known as high spirits, to brawling in public, through to far more serious crimes.

Numerous studies (for example Graham and Bowling, 1995; Flood-Page *et al*, 2000) have established that a small minority of criminals are responsible for a disproportionately large percentage of crime committed. The third group is therefore made up of those who are active, hard-core criminals (C).

There are undoubtedly many factors which govern both membership of and movement between the various categories over time. Some will be of lasting influence, and some will be ephemeral. The aim of the model is to synthesise the key elements which give rise to changes in the relative size of these groups over time, in order to improve our understanding of the dynamics of the behaviour of crime.

An essential part of the model is the influence of social interaction on the behaviour of agents. The importance of this phenomenon is well documented in the literature of criminology and, more generally, in that of sociology itself. Currie (1985), for example, discusses the very wide variations in crime rates which are often observed between the rural and urban sectors of poor economies, attributing in large part the low rates in the rural areas to the community relationships which both foster a sense of belonging and provide *'the setting in which informal social sanctions against aggression and crime can operate effectively'.*

We also introduce potential influences from factors such as demographics, overall social and economic conditions, and the deterrence effects of the criminal justice system, though, to stress again, the literature is very divided as to the relative importance of such variables.

We represent these elements by means of a small number of distinct flows between the three groups of the population. The flow from Non-susceptible to Susceptible is postulated to be due primarily to demographic movements in the population and to social and economic conditions. There is a reverse flow from Susceptible back into Non-susceptible.

There is also a flow out of the Susceptible category and into the Criminal. We assume that part is due to the net effect of social and economic conditions and deterrence. The flow also arises because of social interaction, with susceptible individuals being influenced by the behaviour of those who are already criminals.

Finally, there is a flow from the category of Criminals to the Non-susceptibles. Again, in part a number of people abandon crime in response to the net effect of social and economic conditions and deterrence. In part, the extent of general social disapproval of criminal activity is also postulated to determine this flow.

A formal statement of the initial model

In this section we develop a formal mathematical model of the flows described above. For the most part, in the absence of good evidence to the contrary, we make the simplest possible assumption, namely that the flows into and out of any particular group are proportional to its size. This is perfectly justifiable, since the lack of conclusions in the empirical literature on the impact on crime of the factors listed above means that there is little guidance as to how we should express these factors in a formal, mathematical way.

Consider first of all the flow from the non-susceptible into the susceptible category. The assumption that a constant proportion of the Non-susceptibles flows into the Susceptible category leads to the differential equation for N,

$$dN/dt = -\theta N$$

where θ is a constant which measures the size of the flow. The constant θ must be positive if it is to represent a flow out of N.

There are several other flows into and out of the Non-susceptible category that must be included in the equation for N. There is a flow from the Susceptibles back into the Non-susceptibles which we take to be a constant proportion of S. This gives rise to a term μS in the equation for dN/dt where the constant μ measures the size of the flow, and it too is assumed to be positive.

There is also a flow from the category of Criminals to the Non-susceptibles. The pressure on criminals to drop out of the Criminal category will depend not just (proportionally) upon the net effects of deterrence and social and economic conditions, but also on the extent to which it is the object of social disapproval. For the sake of simplicity, we express social disapproval as a function of the numbers who are in the Non-susceptible category at any point in time. But the impact of disapproval is unlikely to be one of simple proportionality to N. Where N is already high, further increases may well have little effect in influencing the small number who are criminals, whilst for low values of N, in areas where crime has almost become the norm, further declines may also have little effect.

These effects are captured by the function β which is defined by

$$\beta(N) = \beta_1 + \beta_2 / (1 + \exp(-\rho(N - N_C))) \qquad (1)$$

The parameters β_1, β_2 etc. need some explanation. For low values of N, the value of β is approximately β_1 which therefore represents the net effects of deterrence and social and economic conditions. The effects of social disapproval are only felt for sufficiently high values of N when the value of the function is approximately $\beta_1 + \beta_2$. The other parameters N_c and ρ determine respectively the value of N for which social disapproval becomes significant and the sensitivity of disapproval to changes in the value of N. We assume that $\beta(N)$ is the proportion of the Criminals flowing into the Non-susceptibles.

The equation for N can then be written in full,

$$dN/dt = -\theta N + \mu S + \beta(N)C \qquad (2)$$

where C denotes the proportion in the Criminal category.

The equation for the Susceptibles, S, is derived in a similar way. Part of the flow out of the susceptible category into crime itself is due to social and economic conditions, but this is potentially reduced by the impact of deterrence. In addition, we allow for the effect of social interaction. The greater the proportion of a population who are already criminals, the more likely it is that the process of social interaction will encourage others to move from being merely susceptible to crime to committing crime.

The equation for S is therefore given by

$$dS/dt = \theta N - \mu S - \alpha S - \lambda SC \qquad (3)$$

All of the terms in the remaining equation, the equation for C, have already appeared in equations (1) and (2) but they appear here with their signs changed, i.e.

$$dC/dt = \alpha S + \lambda SC - \beta(N)C \qquad (4)$$

Our model is completed by the identity which states that the population, however it is defined, is comprised of the sum of the proportions in the three categories

$$N + S + C = 1 \qquad (5)$$

The proportions could be applied to a population to give the numbers in each of the categories if necessary and the population could be either constant, growing or even falling. Equations (2) to (5) taken together formalise the initial model.

There are several points to bear in mind when discussing the success or plausibility of the model. We are not, to emphasise again, attempting to account for any particular set of observed crime rates, so the standard statistical criteria of econometric work, for example, are not relevant. (It should also be said, however, that given the failure of such work to produce anything like a consensus on the determinants of crime, the apparent rigour of such tests in theory is very much weaker in practice.) We do not attempt to disguise the fact that there are no formal criteria and that ultimately the plausibility of the model is a matter of judgement, but we set out a number of criteria which it should satisfy.

The underlying rationale of the model implies, for example, that a higher value of α, say, or of θ should eventually lead to a higher level of crime — to a greater proportion of the population being in the Criminal category. So it is essential that the model as a whole preserves these qualitative properties which underlie each individual expression in it. It is by no means guaranteed that the system of coupled non-linear differential equations will behave in the way we might expect from each single equation considered in isolation.

A second, very important, criterion of plausibility is that the system should be able to generate a wide range of solutions particularly for the value of C, with respect to relatively small variations in the values of the parameters. An important stylised fact about crime rates is their enormous variation, both over time and, especially, over space. This latter holds not just in terms of international comparisons, but within individual countries. The variation appears to be considerably more substantial than that of factors which might be thought of as possible contributors to the crime rate. So reasonable variations in the values of the parameters should be able to produce a wide range of solutions.

Third, the model must give rise to solutions which are meaningful, in the sense that since N, S and C are proportions of any given population, their values cannot be less than zero, nor greater than one. Therefore, starting from any meaningful set of initial values, the model must not be able to generate solutions which, whilst valid mathematically, nevertheless violate this restriction.

Extensions and modifications of the initial model

An analytical solution can be obtained to a slightly simplified version of equations (2) to (5), in which the parameter β is held constant. By analytical we mean that an exact mathematical expression can be found which gives the solution to the equations. Details of this, and of an extension which allows β to vary endogenously, are given in Ormerod and Campbell (1997), a copy of which has been provided to the Home Office.

It can be shown (Ormerod and Campbell (op.cit.)) that, first, the effect in the overall model of varying the parameters is consistent with what we would expect from the underlying logic of the individual equations. As θ the strength of the flow from N to S, is increased, the equilibrium level of crime, C*, rises. The same is true of the parameters α and λ. If μ the rate at which Susceptibles return to the category of Non-susceptibles, is increased, C* falls. Increasing the effects of social disapproval by increasing the values taken by the function β also leads to a fall in the equilibrium level of crime.

Second, and importantly, the model is consistent with the key stylised fact about crime rates. Namely that the variation across time and space appears to be much more substantial than that of external factors such as social and economic conditions which might be thought of as possible contributors to the crime rate. A very wide range of values for C* can be generated by the model when just one or two of the parameters are varied, sometimes by relatively small amounts.

The initial model is extended in two ways. First, and more importantly, a category for those in prison (P) is introduced. To a substantial extent, the categories of hard-core criminals (C) and those in prison (P) form a self-contained loop in the model, given the high rates of recidivism. However, the introduction of a fourth equation does complicate the analysis of the model even further.

The flow out of P is defined by the simple equation

$$dP/dt = -\varphi_2 P$$

where the parameter φ_2 is the average length of a prison sentence. The flow out of C into P is φ_1C, where the parameter φ_1 depends upon three factors. First, the probability of being convicted per offence committed. Second, the average number of offences committed per period by individuals in the C category, and third, the probability of being sent to prison if convicted.

We introduce the parameter π to measure the division of the flow out of P into the C and N categories. Slightly confusingly, when we analysed the model we defined the proportion moving from P directly back into the C category as $(1 - \pi)$, so the recidivism rate is not π itself but $(1 - \pi)$.

The second change to the initial model is to introduce an explicit term for deterrence into the flow from S back into N. We make this proportionate to the numbers actually in prison, γP. The total number of people convicted is δP, where $\delta > 1$ i.e. people fined, community service etc. So we can think of γ in this model as being $\gamma = \gamma_0 \delta$, where γ_0 is the deterrence parameter, but we simply refer to γ for short in all subsequent references.

The extended model is therefore:

$$dN/dt = -\theta N + \mu S + \beta(N)C + \gamma P + \pi\phi_2 P \qquad (6)$$

$$dS/dt = +\theta N - \mu S - \alpha S - \lambda SC - \gamma P \qquad (7)$$

$$dC/dt = +\alpha S + \lambda SC - \beta(N)C - \phi_1 C + \phi_2 P - \pi\phi_2 P \qquad (8)$$

$$dP/dt = +\phi_1 C - \phi_2 P \qquad (9)$$

and the corresponding normalisation condition becomes $N + S + C + P = 1$.

The system (6) to (9), along with the expression for $\beta(N)$ given by (4), has been solved numerically (using a 4th order Runge-Kutta algorithm) and equilibrium values of the system (N^*, S^*, C^*, P^*) obtained for different values of the parameters of the system.

We did consider whether to focus on this non-linear differential equation (NLD) approach or whether to adopt an explicit interacting agent based approach. Our feeling was, on balance, to investigate in this project the potential of the NLD approach to increase understanding of crime. There are several reasons for this:

i) both the NLD and the agent-based approach allow explicitly for social interaction between agents

ii) under certain assumptions, the two approaches can be regarded as equivalent (rather, the stochastic version of the deterministic NLD system set out above can be)

iii) the advantages of the agent-based approach are that

a) it permits a more explicit representation of difference in behaviour between individual agents (which is implicit in the NLD approach)

b) agents can be connected on a variety of topologies

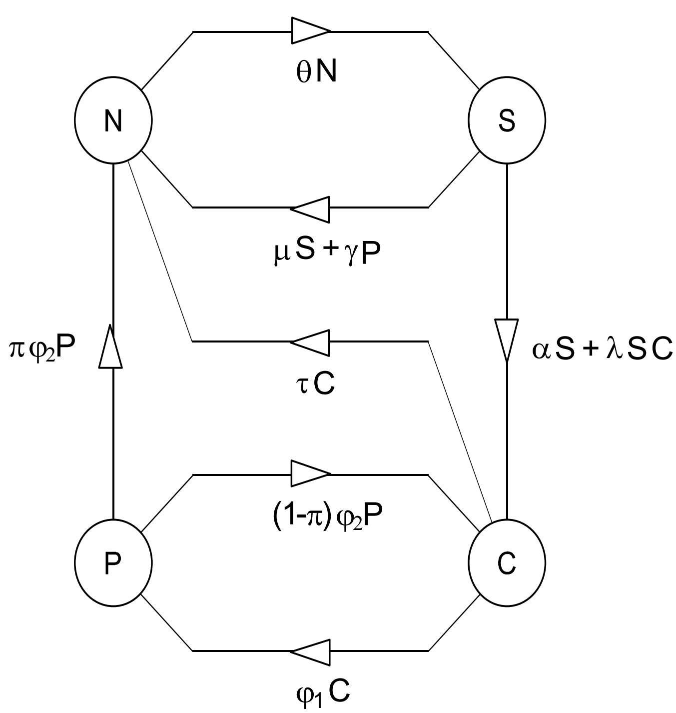
iv) the advantage of the NLD approach is very practical. Innovative models such as the NLD one above or an agent-based approach give rise to complicated behaviour of the system as a whole. With the NLD approach, we are able to understand this with an analytical solution (which still requires considerable effort to explore thoroughly in terms of different combinations of values of the parameters). With the agent-based approach, we have to search this complex phase space by simulation.

Before proceeding to a calibration of the model to actual data, there is one further modification which we made. It has not been possible to obtain an analytical solution to the system described by equations (5) to (9). We therefore approximated the non-linear β by a linear function, details of which are set out in Appendix 1. Non-linearity does, however, still feature in the model in the flow from S to C.

By using the analytical solution, we are readily able to compare the equilibrium values of N, S, C and P which emerge for different values of the parameters. This can also be done numerically, but the analytical approach enables us to have more confidence that the potential solution space, particularly regarding the local stability of the equations, has been thoroughly explored. The analytical solution to this modified version of (5) to (9) is given in Appendix 2.

Schematically, the model can be represented in Figure 1 below.

Figure 1: Schematic of flows in crime model



3. Calibration of the model to property crime data

The late 1990s

In this section, we discuss plausible ranges of values for the parameters of the model. From within these ranges, we are able to obtain a solution to the model which gives a level of property crime (i.e. burglary) similar to that which was obtained in the UK in the late 1990s.

The choice of the base population is obviously an important issue. In this initial stage of the model development, our concern was to produce a calibration which is reasonably realistic for current levels of property crime. This is clearly subject to refinement, but the aim was to produce a reasonable base solution against which to work.

Property crime is committed overwhelmingly by young men. Within this group, there is a further strong concentration amongst the poorer and less educated sections. The total male population of the UK aged between 15 and 30 is around six million, with 90 per cent of this figure in England and Wales alone. Of these, very few of the economically and educationally advantaged will ever commit property crime. As a working hypothesis, we calibrate our model using a base population from the bottom quartile (i.e. the poorest 25 per cent) of this male age group, in other words, 1.5 million people.

We make the assumption that those in the S category commit on average one crime per year and that criminals commit ten crimes per year. In recent years there have been around 600,000 domestic burglaries per year and currently there are around 8,000 men in prison for committing burglaries. This means that we should be looking for equilibrium proportions that break down our population in the way shown in Table 1:

Table 1: Breakdown of crimes committed in the late 1990s

Category	%	Numbers (1000s)	Crimes per person per year	Crimes per year (1000's)
Non susceptible	82	1,230	0	0
Susceptible	15	225	1	225
Criminal	2.5	37.5	10	375
Prisoner	0.5	7.5	0	0
Total	**100**	**1,500**		**600**

The interpretation of the parameters in the model is in some cases straightforward, whilst others are attempting to summarise in a single measure a more complex range of causal factors.

The parameter φ_2, for example, is the proportional flow out of the prison category. The rate of change of the level of P depends negatively upon P itself multiplied by φ_2. The parameter shows the proportion of prisoners that are released each year. From the data in Deadman (2000), the average prison sentence for burglary has varied over time between around one year and 18 months. A value of φ_2 of 0.6 implies that 60 per cent of prisoners are released each year, and we choose this as the base value.

The parameter φ_1, whilst still quite simple to interpret is nevertheless more complex than φ_2. It shows the proportion of the C category who flow into P in any given year. This is a combination of the probability of being convicted per burglary carried out and the probability of being sent to prison if convicted. Again from Deadman, we see that convictions as a proportion of recorded burglaries is just under four per cent in the late 1990s, and the proportion of those convicted sent to prison is between 20 and 25 per cent. By assumption, individuals in the C category carry out ten crimes a year, so the probability of an individual being convicted in a given year is around 35 per cent. This implies a value of φ_1 of around 0.1.

The parameter π represents the proportion of those leaving prison who reform. In terms of the model, they move into the N category. The remainder go back into the C category. The initial value of this parameter is set at 0.25, implying that 75 per cent of released prisoners reoffend rapidly.

Direct information with which to calibrate the other parameters is less readily available, though we can make reasonable deductions. Consider, for example, the parameters ϑ and μ. These regulate the flows between the N and S categories. The parameter ϑ is the proportional flow from N to S and μ the proportional flow back from S to N. These parameters summarise a complex range of factors. For example, the economic circumstances which motivate people in moving between these two categories, and the availability of goods to burgle. There is also the potential factor of social influences, in other words the degree of disapproval which attaches itself to occasional crime. We could complicate the model still further by separating out this latter factor and making it a non-linear function of N and S. Finally, there will be circumstances particular to individuals, such as getting married, which was traditionally an important influence on extracting young men from lives of occasional crime.

The population in the model is defined as being between the ages of 15 and 30. Imagine an individual coming into the model, as it were, into the N category at the age of 15. Each year he remains in the system, the probability in any given year of moving into the S category is given by ϑ. An individual may stay in the N category throughout the entire length of time he is in the model. He may move into S for just one year, and exit back to N with probability μ. Once back in N, he may re-enter S, and so on. By focusing solely on these flows between N and S as a sub-loop of the model, we can obtain an estimate of the probability of an individual arriving at age 30 of not having committed a crime. In other words, of remaining in the N category all the time. The details of this analysis are set out in Appendix 3. Choosing values of ϑ and μ of 0.08 and 0.3, respectively, implies that, as an approximation, a young man in our population has a 45 per cent chance of not committing a crime during his time in the at risk age group. This seems a reasonable base value for the initial calibration of the model.

The parameter γ is another influence on the flow from S to N, but in this case it is proportional to P, the prison population. It represents the deterrence effect of prison. We refer in this note simply to γ. However, the total number of people convicted is δP, where $\delta > 1$ i.e. people are fined, given community service and so forth, rather than being sent to prison. So we can think of γ in our model as being $\gamma = \gamma_0\delta$, where γ_0 is the deterrence parameter.

Estimates of the impact of deterrence in the crime literature vary substantially, and it is not possible to draw on this to determine the value of γ. In the initial calibration, we set γ equal to 2. This may seem high, but the numbers in P are very small compared to the numbers in S. To anticipate the results of this calibration exercise, it is implied that P is around 3 per cent of S. With γ equal to 2, some 6 per cent of those in S are assumed to flow back into N each year because of the impact of deterrence.

The parameter α captures the net effects of several different factors: deterrence; general economic conditions on the decision to move from being an occasional to a hard-core criminal; the availability of goods to burgle. It measures the proportional flow from S to C for these reasons. The fact that a substantial proportion of total crime is committed by a very small proportion of the population implies that C is small relative to S. This in turn implies that α is small. Otherwise, given the values of the parameters which determine the flow out of the C category (φ_1 and τ), the numbers in the C category would quickly become too large. We therefore choose 0.015 as the initial value for α.

The parameter λ represents the flow from S to C due to social interaction. In other words, the influence of those who are already criminals in the decision to move from the S to the C category. The parameters τ_1 and τ_2 represent the drop-out rate, as it were, from hard-core criminality and the decision to go straight. A whole variety of factors will influence the value of τ_1, such as age and getting married. This is simply a constant flow out of C. The parameter τ_2 represents the flow out of C due to social pressure, which is proxied by the numbers in the N category. Other things being equal, the greater the proportion of the population who are not committing crimes, the greater the social pressure on criminals to reform.

Given the values already chosen for the other parameters in the model, we need to choose a set of values for λ and the τ parameters which gives a model solution consistent with the number of recorded burglaries observed in the late 1990s. There is clearly far less evidence available on these parameters than on others in the model. But a satisfactory set of values is given by $\lambda = 1$, $\tau_1 = 0.12$ and $\tau_2 = 0.13$.

This set of parameters yields a solution to the model which generates 600,000 burglaries per year. The equilibrium levels for the values of the categories are:

N = 81.5% S = 15.7% C = 2.4% P = 0.40%.

The number of those in prison for burglary which this implies is slightly lower than that which actually obtains. However, as an initial calibration, this seems reasonable.

The early 1950s

We can go on to consider, in less detail, an initial calibration of the model for the number of burglaries committed in 1950. An important point to consider is the number of burglaries individuals commit each year. It seems reasonable to continue with the assumption that those in the S category commit one a year. For more frequent offenders, the question arises of what we might term the productivity of burglars. Like most other economic activities, burglary must be subject to productivity gains. It will become easier to commit a given number over time. To the extent that we can view burglary as a service sector activity, productivity gains may not be as large as those observed in manufacturing. There will also be interplay with the response of householders to security, which will further limit productivity growth.

The question as to whether productivity gains imply less or more burglaries per head committed over time each year by hard core criminals does not have an unequivocal answer. To achieve a given level of real income, as burglary becomes easier, fewer might be committed by each individual each year. Against this, the reward for any given level of effort becomes greater, incentivising an increase in per capita terms. For the purposes of calibrating the model against the levels observed in the early 1950s, we have assumed that this latter effect has predominated over time, although it is straightforward to re-calibrate with a different assumption. As an initial solution for the model, we assume that each hard core criminal committed 7.5 burglaries per year in the early 1950s compared to ten in the late 1990s.

On this basis, a plausible breakdown of the categories and the levels of burglaries associated with them is set out in Table 2.

Table 2: Breakdown of crimes committed in the 1950s

Category	%	Numbers (1000s)	Crimes per person per year	Crimes per year (1000s)
Non susceptible	97.55	1,170.6	0	0
Susceptible	2	24	1	24
Criminal	0.4	4.8	7.5	36
Prisoner	0.05	0.6	0	0
Total	**100**	**1,200**		**60**

First of all, we can consider the impact of varying individual parameters from their values in the 1990s calibration, and see how far each goes in producing solutions similar to those observed in Table 2. For reference, the equilibrium values for the various categories of the model in the late 1990s calibration are:

N = 81.5%   S = 15.7%   C = 2.4%   P = 0.40%.

The scope for changing some of the parameters is very limited. For example, it is not unreasonable to retain the value of 0.25 for π, the parameter which measures recidivism. In any event, changing the value of π within a plausible range has little impact on the level of crime implied by the model. If more of those in prison are reformed, so that as many as 40 per cent go from the P to the N category, the equilibrium solution of the model gives S = 16.1% and C = 2.2%, very little different to those estimated for the late 1990s.

Although the average prison sentence has varied over time, in the early 1950s it was only slightly higher than it was in the late 1990s. Instead of 0.6 for φ_2, we set this parameter at 0.55. Perhaps surprisingly, one of the components of φ_2 has not varied between the two periods. The probability of being sent to prison if convicted is very similar. However, the probability of being caught was much higher in the early 1950s, at around 0.20 per burglary compared to 0.04 in the late 1990s. Against this, we have assumed that the average person in the C category committed only 7.5 rather than ten burglaries per year. The net effect is to raise φ_1 to around 0.16.

The net effect of these two parameter changes on the 1990s calibration is small. The value of S falls from 15.7 per cent to 15.3 per cent, and of C from 2.4 per cent to 2.2 per cent. If these had been the only changes which had taken place, annual burglary levels in the early 1950s would have been some 380,000.

The final aspect of the criminal justice system in the model is γ, the impact of deterrence. A reasonable assumption to make is that this has fallen over time, since the probability of being convicted per burglary has fallen to a fifth of its early 1950s value. If we set γ at a value of ten, five times its value in the late 1990s model calibration, the levels of S and C become 11.5 per cent and 1.2 per cent respectively. This is quite a substantial change, though it is based on an assumption that the fall in the probability of conviction over time is entirely reflected in a fall in deterrence effects. Certainly, it is difficult to argue that γ should be more than ten in the early 1950s.

We now consider the section of the model describing the flows between N and S, and the parameters ϑ and μ. In the late 1990s calibration, the choice of 0.08 and 0.3, respectively, for these parameters implies (as an approximation) that 45 per cent of the relevant population group never offend. Casual empiricism suggests that this proportion was certainly higher 40 or 50 years ago. As an illustration, suppose 75 per cent never offended. This value can be obtained by ϑ = 0.02 and μ = 0.075. These reductions, when combined with the values of the other parameters in the 1990s calibration, give levels of S and C of 11.6 per cent and 1.2 per cent respectively.

The remaining parameters, α, λ and τ, are concerned with the categories S and C. In the 1990s calibration, α is already at the arithmetically low value of 0.015. The stock of durable goods available to steal increased by a factor of three in the second half of the 20th century. If this is reflected on a one-for-one basis in α, the implied value of α in the early 1950s calibration is 0.005. Further, the proportion of income accruing to the bottom quintile has fallen over time. Marris (2000) estimates this at between nine and ten per cent in the early 1950s, and some seven per cent in the late 1990s. Again, making the heroic assumption that this translates on a one-for-one basis in α, the 1950s value of α falls even further to (7/9.5)*0.005 = 0.004. This value for α, along with the 1990s values for other parameters, gives values for S and C of 18.5 per cent and 1.1 per cent respectively.

The explicit social value parameters, λ and τ_1, are, as already noted, less easy to characterise. In some ways their values can be seen as absorbing the slack in the system. Given a choice of values for parameters on which we do have (varying amounts of) information, setting λ and τ_1 can help calibrate the model around a particular solution.

The impact of λ is already to some extent self-reinforcing, for it enters into the equation for the flow out of S in the form λSC. In other words, if changes in other parameters lead to lower values of C, even with λ constant the impact of social interaction will diminish the flow out of S into C. In the original theoretical specification of the model, the τ_1 also entered in a non-linear way (denoted as $\beta(N)$ in the original model). However, as noted above we could not then obtain an analytical solution of the model, and a simplification was made. There is still a self-reinforcing element, because the flow out of C into N depends upon $\tau_2 N$, so that the lower are S and C due to changing other parameters, the higher is N, and in consequence the lower still is C.

That said, it is difficult to argue with any degree of precision for changes in these parameters over time. By way of illustration, however, if λ were reduced to 0.5 instead of 1, S changes to 17.2 per cent from 15.7 per cent, and C from 2.4 per cent to 1.6 per cent, and if τ_2 were increased to 0.2 from 0.13, S and C become 16.8 per cent and 1.8 per cent respectively.

Considering these effects in their entirety, the model clearly can generate solutions which span the 1950s level of crime, certainly if taken in conjunction with each other.

4. Calibration to data on violence against the person

Background information in the calibration of the late 1990s model

As with property crime, violent crime is committed overwhelmingly by young men. Within this group, there is a further strong concentration amongst the poorer and less educated sections. The total male population of the UK aged between 15 and 30 is around six million, with 90 per cent of this figure in England and Wales. Of these, very few of the economically and educationally advantaged will ever commit violent crime. As a working hypothesis, we calibrate our model using a base population from the bottom quartile of this male age group, in other words, 1.5 million people.

The Cambridge Study of Delinquent Development is useful in calibrating the application of the model to violent crime. The study started in 1961 with a survey of 411 boys, ages eight to nine, living in working class neighbourhoods in London. All the boys lived within a one-mile radius of the research office that had been established for the study. The population of this area is very similar to the characteristics of our overall population, namely poor and relatively unskilled and uneducated.

By 1985, the group had a total of 82 violent crime convictions from 51 of the men. The total number of crimes committed was 683 from 153 men, but half of these were committed by only 22 of the men. This latter group is 5 per cent of the total population in the sample of 411. Making the assumption that the distribution of violent crimes is the same as the distribution of all crimes in the study, we can split the 51 violent criminals into two groups – occasional and hard core. Of the 153 men committing any crime, 22 committed one-half, which implies that eight out of the 51 violent criminals were hard core.

In order to see how many recorded crimes per year the study implies in the overall population, we need to divide by 24 (the number of years over which the study took place), and multiply by the scalar difference in the population sizes (the average annual population of young men 1961 to 1985 is similar to that of the late 1990s). This will give us the equivalent number of convictions, which then needs to be grossed up to the number of crimes.

The conviction rate for the late 1990s was around 1:7, though from the more detailed data available to us on property crime, we know that the probability of conviction has fallen over time. With burglary, property convictions as a proportion of recorded crimes was far higher than for violent crime in the late 1990s, being around 1:5, and ranging between 1:5 and 1:10 over the period of the Cambridge study. The ratio of 1:5 was also obtained in the 1950s, suggesting some sort of lower practical bound. We therefore assume a proportion of 1:5 for violent crime in the period of the Cambridge study.

With this assumption, we obtain an overall figure of 62,500 recorded violent crimes a year, which compare with the actual recorded annual number averaged over the period of the Cambridge Study (1961 to 1985) of around 65,000 per year. The Cambridge Study was, of course, designed to be representative, and we have found an implied overall level of convictions which is consistent with this.

We now apply the late 1990s ratio of convictions to recorded crimes of 1:7 to the information in the Cambridge study, and we obtain what recorded violent crime would have been in the late

1990s, along with the breakdown between our model categories and their annual crime rates, if this had been the only factor which had changed. This is summarised in Table 3.

Table 3: Populations of Susceptible and Criminal categories and their annual crime rates, implied by the Cambridge study with late 1990s conviction: recorded crime ratio

Category	Susceptible	Criminal	Total
Study population	43	8	51
Study crimes	12	12	24
Crime rate	0.28	1.5	0.47
Model population (k)	157	29	186
Implied model crimes (k)	44	43	87

Notes:
Study population refers to the Cambridge study discussed above
Model population refers to the overall violent crime population, total of 1.5 million
Crime Rate = Study crimes over Study population
Model population = Study population x population scaling constant
Implied model crimes = Model population x Crime rate

In terms of percentage of the population in the two categories, the figure for the S category is 10.5 per cent and that for the C category 1.9 per cent.

The main difference between the property and violent crime calibrations is the much lower *rate* of recorded crime in the case of violence. The low rate of recorded crime, particularly amongst the Susceptible category, is supported qualitatively by a wide range of studies of the behaviour of boys from deprived backgrounds. In his classic study of a violent young Glasgow gang in the late 1960s, for example, James Patrick (*A Glasgow Gang Observed*) notes that his experience was identical to that reported in the American literature, namely that an important activity for gang members was to be seen to put themselves into situations where fighting looked likely, without actually having to fight. Of course, violence against the person is by no means solely a gang activity, but we can think of this mentality applying more widely amongst the social groups of young men vulnerable to committing violent crime.

Benchmark calibration of the late 1990s model

The single most striking feature of Table 3 is of course the fact that the actual level of recorded violent crime in the late 1990s was substantially larger than this figure of 87,000. Comparisons are complicated by the extensive definitional changes introduced in 1998, but the average in the late 1990s on the same basis was around 265,000. This implies that either or both of the crime rates or the crime populations have changed since the period of the Cambridge study. Of course, the difference between the two periods is by no means as large if British Crime Survey data is used. As Marris indicates in his review, between 1981 and 1999, for example, recorded violent crime grew by some 165 per cent, but the rise in the BCS victim-reported data was only 50 per cent . Even so, the increase was still substantial. We proceed with the calibration using the recorded data, since this is what we were asked to do by the Home Office.

Suppose that we make the most straightforward assumption, namely that the increase is equally divided between the two categories and examine the implications. It should be stressed immediately that the available evidence for the late 1990s suggests strongly that this is *not* what has actually happened. Most of the increases in crime appears to have arisen from changes in the C category. But is useful to carry out this across-the-board uprating in order to compare the outcome with what the actual evidence implies.

Multiplying the annual crime rates for the S and C categories by 1.75 gives rise to a $1.75^2 = 3.05$ times rise in the overall recorded VATP. We then have a population breakdown for the model as shown in Table 4.

Table 4: Populations of Susceptible and Criminal categories and their annual crime rates for the late 1990s implied by across-the-board increases on the Cambridge study figures

Category	Susceptible	Criminal	Total
Model population (k)	157 x 1.75 = 275	29 x 1.75 = 50.75	
Crime rate	0.28 x 1.75 = 0.49	1.5 x 1.75= 2.63	
Implied model crimes (k)	135	133	**268**

In other words, this implies a solution to the model for the late 1990s in which some 18.3 per cent of the relevant population are in the Susceptible category, compared to the 10.5 per cent average over the 1961 to 1985 period, and 3.4 per cent in the hard core Criminal category compared to 1.9 per cent. Actual data is available on the number in prison for violent crime against the person, a figure of just over 10,000, putting the P category at 0.67 per cent of the relevant population.

Implications for violent crime in the late 1990s

Information on violent crime with which to calibrate the model is not as extensive as with property crime, particularly over time. But a key section of the model where we do have further information is in the flows between the C and P categories, which has important implications for how the distribution of crime has evolved since the period of the Cambridge study.

For the record, we keep the assumption of a high rate of recidivism and keep π at 0.25 as in the property calibration.

The value for φ_2, the flow from P to C, is determined by looking at the length of the average prison sentence. For the late 1990s this was approximately one and a half years. The value of φ_2 equals the inverse of the average sentence, and the value used is 0.65.

In choosing φ_1, the flow from C to P, we know the number of convictions per offence reported, and the proportion of those convicted being sent to prison. This implies that the probability of being sent to prison per reported offence is only some 0.03 i.e. three per cent. The value of φ_1 can be readily calculated using simple probability theory. The parameter φ_1 measures the probability of moving from C to P in any given year, and is identical to one minus the probability of *not* being sent to prison in any given year. This latter is in turn defined as $[1 - (1 - 0.03)^n]$ where n is the average number of offences carried out each year by an individual in the C category (these logical steps are also set out in Appendix 4).

We know P, the percentage of the relevant population in prison, φ_2, and the probability of being sent to prison per offence. We do not know C, the percentage of the relevant population in the hard core criminal category, and we do not know n, the number of recorded crimes committed each year by these individuals. But we also have one further piece of information from the analytical solution to the equations of the model, namely that in equilibrium, the relationship between C and P is as follows:

$$C = (\varphi_2/\varphi_1)\, P$$

Of course, as we noted in the section on the property crime calibration, recorded crime data do not suggest that the system is actually in equilibrium. The rate of change of recorded crime itself changes from year to year. However, the fluctuations are not too extreme, and as an approximation we might think of the system as being in the neighbourhood of equilibrium at any point in time, and changes in crime are brought about by changes over time in parameter values.

Above, we applied across the board increases in both the numbers in the S and C categories and their annual crime rates to the Cambridge study figures, and obtained an annual crime rate for the C category of 2.63 crimes per year.

Using this value of 2.63 for n gives a value of φ_1 of 0.08, which implies that the C category comprises just under 5.5 per cent of the total population, or some 82,000 people. This is substantially larger (in percentage terms) than the 51,000 implied by across-the-board increases in the Cambridge study figures. And in turn it suggests that some 215,000 out of the total of 275,000 crimes were committed by the hard core category in the late 1990s, almost 80 per cent compared to the 50 per cent observed in the 1960s and 1970s.

An alternative way of approaching this issue is to take the percentage in the C category implied by across-the-board increases and, using the same information as above, work out the implied number of crimes committed per year by hard core criminals. Without going through the calculations in detail, this implies an average of 4.3 crimes per year, compared to the across-the-board number of 2.63. But this also implies that the number of crimes committed by those in the C category is 230,000.

In other words, the clear implication is that most of the increase in violent crime which has taken place since the period of the Cambridge study has been the responsibility of individuals in the hard core, C, category.

Appendix 4 sets out the way in which information is used to calculate C and n, and the annual number of crimes which the various values imply. As Appendix 4 shows, the number of crimes committed by those in the C category is very insensitive to the values of C and n which are chosen. Intuitively this is because φ_1 increases at a rate that is roughly linear with respect to the average number of crimes carried out per year per individual (for low values of this latter number). Therefore, the number of crimes, which is proportional to the individual crime rate divided by φ_1, remains relatively constant.

Obviously, a mixture of changes in both the numbers in the categories and the annual rates of crime could lead to an appropriate solution. We can obtain a further perspective on this by calibrating the model for the period covered by the Cambridge crime study, and then using information on possible changes in parameters which might have taken place since then.

Calibrating on the period of the Cambridge study

The Cambridge study is over the period 1961 to 1985, so we consider the level of violent crime which obtained in the mid/late 1970s (given that the boys were only 8 or 9 when the study started, we take a point somewhat beyond the mid-point of the period). The parameters θ and μ control the bulk of the flows between the N and S categories. The parameter θ is the proportional flow from N to S and μ the proportional flow back from S to N. These parameters cover a wide range of factors, including the social disapproval of committing violent crimes, and the social and economic circumstances of low skilled youths.

In the Cambridge study, 51 out of 411 men had convictions for violent crime, or 12.5 per cent. As described in the property calibration, and repeated here in detail in Appendix 3, the values of ϑ and μ determine the proportion of the population which either remains in N all the time, or has been in S at some time but is now back in N, or is now in S. The choice of $\vartheta = 0.06$ and $\mu = 0.4$ implies that the latter two make up 47 per cent of the population. In other words, an individual has a 53 per cent chance of remaining not susceptible to crime as he passes from the ages of 15 to 30. The number of crimes committed per year by members of the S category is around 0.3, and the average amount of time spent in S, it can be shown, is approximately three years. So, for an individual in S the probability of actually committing a crime in any given year is 0.3, and of not committing it is 0.7. Again, simple probability theory tells us that the chance of an individual who is either in or has been in S actually committing a crime is approximately two-thirds. If 47 per cent of the population are either in or have been in S, this implies that just over 30 per cent have actually committed a violent crime. Comparing this to the figure of 12.5 per cent who have actually been convicted, it seems reasonable.

We assume that the parameter describing the recidivism rate ($1-\pi$) was the same in the mid-1970s as it is today, at 0.75. Further, we assume the average jail sentence was the same, giving $\varphi_2 = 0.65$. There is some evidence from the property data that sentences were somewhat lower 25 years ago, but we have no direct evidence on violent crime. The value of φ_1 implied by the data on annual crime rates per person in the C category and by our various assumptions is 0.064. In other words, there was a 6.4 per cent chance in any one year of an individual in the C category being sent to prison.

In the absence of other information, we begin by setting the other parameters to their values calibrated in the property model. The reasons were set out in detail in the property model paper.

These values result in S being 11.4 per cent of the population and C being 1.2. The actual values of S and C were 10.5 and 1.9, respectively. A simple change in the value of α from 0.015 to 0.025 gives a solution of S = 10.6 and C = 1.8. This, therefore, seems a reasonable calibration for the mid/late 1970s model.

Calibration of the late 1990s

The survey by Marris (2000) identifies from the empirical literature three key potential influences on violent crime:

- real income
- alcohol consumption
- inequality.

However, Marris criticises strongly the findings on real income, mainly because their lack of theoretical underpinnings, which leaves us to consider alcohol and inequality. Marris also identifies an additional factor, namely social values, and cites a small amount of recent work which provides support for this view. The model, of course, includes the parameter λ which expresses the impact of social interaction in the decision of individuals

We now consider how these have moved over the past 20 years and their implications for various parameters. Once we have done this, we will be able to narrow down even further the choice of φ_1 required to calibrate the model for the late 1990s and, by implication, the choice of n, the number of violent crimes committed per year by individuals in the hard core C category.

The first thing to notice is that the amount of alcohol consumed in the UK, in alcohol litres, is very similar in both the mid 1970s and the late 1990s. The vast majority of this is beer, still well over 80 per cent of the total, which has actually fallen slightly over the past 20 years or so, with increases in wine replacing it. It is therefore hard to argue that the amount of alcohol consumed per se has had any impact on the amount of violent crime. Both beer and total alcohol consumption are plotted over time in Appendix 5. Looking in more detail at patterns within alcohol consumption, the data suggests quite clearly that they are stable over time.[2] There are shifts towards greater consumption by women, and increases in drinking by the early and mid-teens, and evidence of some rise in the numbers who drink heavily. But these are all second-order changes against a stable overall background.

In contrast, the degree of inequality has altered sharply. The percentage of income taken by the bottom quintile has fallen from some 10.5 per cent in the mid 1970s to only seven per cent in the late 1990s. Assuming this feeds through on a one-for-one basis onto the relevant parameters, ϑ and α, the respective values implied for the late 1990s are 0.09 and 0.0375 instead of 0.06 and 0.025.

These values of the parameters give values of S and C of 12.9 and 4.0 respectively. However, with φ_1 at its 1970s value, the value for P is only 0.39, just over one-half of its known value of 0.67. Appendix 4 shows that a C value of 4.0 implies a φ_1 of 0.11, and using this gives values for S, C and P of 12.5, 3.45 and 0.58 respectively. A further iteration, as it were, using the value of φ_1 (0.127) implied by this new value of C gives S = 12.4, C = 3.3 and P = 0.64, very similar to the actual for the number in prison.

The implied average number of crimes committed by individuals in the C category is some 4.3, with the C category committing 215,000 crimes in total. The total number in the S category implies that the annual rate of crime for these individuals has remained very low, at between a quarter and a third per year, implying around 50,000 crimes committed by these individuals. In total, violent crime is calibrated at 265,000 a year, very close to the (estimated) actual values at the end of the 1990s. In other words, the model suggests that it is possible to account for the increase in violent crime since the 1970s by movements in the degree of inequality.

This is not to say, of course, that individuals are necessarily motivated directly by inequality. But rather, as the survey by Marris (2000) points out, that it is the overall circumstances associated with poverty which can lead to an increase in violent crime.

For interest, an initial sketch can be made of a calibration against the sorts of level of violent

[2]Data is available in the Department of Health Statistical Bulletin 24October 1999, *Statistics on Alcohol: 1976 Onwards*

crime implied by the increases in the BCS data over the 1981 to 1999 period. The calibration of the model to the period of the Cambridge Study gives a level of recorded crime of 62,500, a figure which obtained around 1973/74. Between then and 1981, the first year of the BCS, there was an increase of some 60 per cent in recorded crime. As noted above, over the 1981 to 1999 period, BCS data show a 50 per cent increase. Heroically splicing the 1973/74 to 1981 recorded percentage increase and the 1981 to 1999 BCS increase gives a total increase between the early 1970s and 1999 of 140 per cent. This implies what we term an adjusted level of recorded crime of some 150,000 in 1999, compared to the 265,000 used above in the calibration.

Suppose we retain the same assumptions as in the 1970s calibration, including the annual number of crimes committed per individual in the S and C categories and the conviction rate. We then adjust α and ϑ as above, to 0.0375 and 0.09 respectively. The resulting level of crime implied is 144,000, close to the actual adjusted level.

5. Policy implications of the model

We can use the solutions of the model calibrated against the late 1990s values for both property and violent crime to assess the impact of the various policy factors which are included in the model. The parameter φ_2, for example, gives the average length of prison sentences, and by varying its value we can observe how the numbers in the S and C categories of the model change. It must be emphasised that we are not saying how changes in the relevant parameters of the model might be brought about, but are examining the consequences if the changes did take place.

We consider a number of policy changes, both singly and in various combinations. Individual impacts are set out for doubling the length of prison sentences, for reducing recidivism from the assumed 75 per cent to 25 per cent, and for doubling the probability of criminals being sent to prison. For a given number of crimes per individual, this latter could be brought about by either by an increase in police efficiency in catching criminals, or by a change in sentencing policy so that more of those convicted are sent to prison. We also present results for changes in the punitive aspects of the criminal justice system combined with an increase in the deterrence effect.

Table 5: Model estimates of the reductions in crimes resulting from policy changes

Policy	**Percentage fall in burglary**	**Percentage fall in violence against the person**
Double length of prison sentences	10	15
Reduce recidivism from 75 to 25 per cent	10	15
Double the probability of criminals being sent to prison	15	20
Double length of prison sentences and deterrence effect doubles	30	35
Double length of prison sentences *and* reduce recidivism from 75 to 25 per cent	20	25
Double length of prison sentences *and* double the probability of criminals being sent to prison	30	40
Double length of prison sentences *and* double the probability of criminals being sent to prison *and* deterrence effect doubles	45	55

It is important to note that the changes in the equilibrium levels of crime due to the changes are rounded to the nearest five per cent, to avoid the impression of spurious accuracy.

Even the quite dramatic changes posited in the criminal justice system have, in isolation, a relatively small impact on crime. Doubling prison sentences, for example, reduces burglaries by 10 per cent and violent crime by 15 per cent. Completely altering recidivism rates has a very similar impact.

The reason for this is that the process of *creating* criminals is not directly affected by these measures. The criminal justice system per se operates on those already in the C and P categories, and does not impact directly on the flows from N to S and into C.

The effect of influencing the flows between the N and S categories can be seen when the impact of deterrence is changed. This is assumed to influence young men in the occasional crime category, S, and to persuade them to move back into the N category rather than remain in S, with a risk of graduating into the C category. The empirical crime literature, as noted above, gives no firm indication of the impact of deterrence, although on balance it might be thought to be more likely to arise from changes in the probability of being caught rather than by sentencing policy.

Table 5 shows that *if* deterrence changes along with changes in the punitive aspect of the criminal justice system, the impact on crime is much greater than it is with changes in the latter factors alone.

For interest, we report finally the effect in the model of varying one of the social interaction parameters, λ. This is a factor in the flow from S to C, in the form λSC. Increasing, for example, the value of λ implies that it becomes more fashionable, as it were, to move from the S to the C category.

Doubling the value of λ leads to a 60 per cent increase in property crime and a 45 per cent increase in violent crime from the levels of the late 1990s. The non-linearity of the way in which λ appears in the model is reflected in the fact that halving its value leads to falls of only 20 and 15 per cent respectively.

6. Conclusion

The purpose of this project is to investigate an innovative approach to the evolution of crime over time using non-linear differential equations.

The intellectual origins of the particular method adopted here are in mathematical biology and, specifically, in how epidemics either spread or are contained over time. A population is divided conceptually into four categories, those not susceptible to commit crime, those committing occasional crimes, hard-core criminals and those in prison. The distinction between occasional and hard-core criminals, for example, is very well documented, with a small minority being responsible for a disproportionately large amount of total crime.

Flows between these four stocks of people are described by differential equations. The choice of parameters with which to describe these flows is drawn from evidence in the criminological literature.

Choosing as the relevant population the number of young, relatively unskilled men, the model has been calibrated successfully to the levels of burglary and violence against the person which prevailed in England and Wales in the late 1990s.

The study shows that the development of models along these lines is certainly feasible.

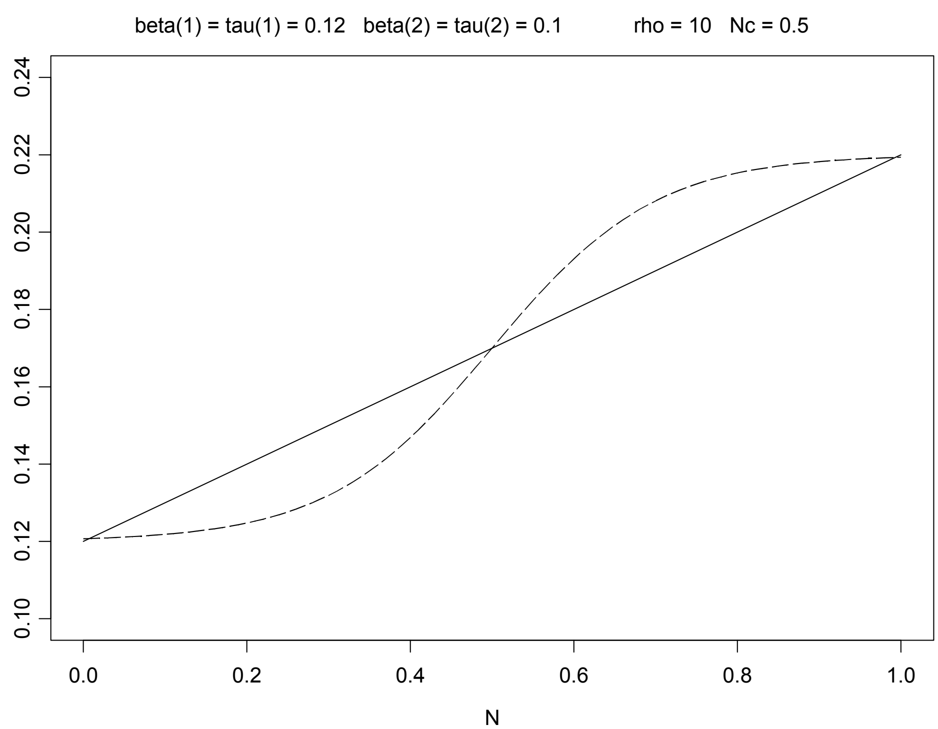
The model allows the impact of policy changes to be assessed. For example, we investigate the effect of doubling prison sentences and of changing the rate of recidivism. An important implication of the model is that, whilst such changes do have impacts on crime, these are not large, because the process of creating criminals is not directly affected by these measures.

Appendix 1. Simplifying the β function

The flow from C to N, previously βC, has been replaced by τC. The original sigmoid Beta function was dependent on four external parameters, β_1, β_2, ρ and N_c. This led to difficulty in obtaining an analytical solution of the model. Thus, the replacement linear function τ is only dependent on two external parameters, τ_1 and τ_2:

$$\tau(N) = \tau_1 + \tau_2 N$$

Figure 2: Comparison of τ and β functions



τ_1 and τ_2 are roughly equivalent to β_1 and β_2 respectively:

$$\tau(0) = \tau_1 \qquad \beta(0) \approx \beta_1$$
$$\tau(1) = \tau_1 + \tau_2 \qquad \beta(1) \approx \beta_1 + \beta_2$$

Using this function does not exclude the possibility of returning to a sigmoid function at a later date.

Appendix 2. The solution and stability of the non-linear differential equations of the model

The final set of equations used is:

$$\frac{\partial N}{\partial t} = -\theta N + \mu S + \tau(N)C + \pi\varphi_2 P + \gamma P \quad (1)$$

$$\frac{\partial S}{\partial t} = \theta N - \mu S - \lambda SC - \gamma P - \alpha S \quad (2)$$

$$\frac{\partial C}{\partial t} = \alpha S + \lambda SC - \tau(N)C - \varphi_1 C + \varphi_2 P - \pi\varphi_2 P \quad (3)$$

$$\frac{\partial P}{\partial t} = \varphi_1 C - \varphi_2 P \quad (4)$$

$$N + S + C + P = 1 \quad (5)$$

where $\tau(N) = A + BN$ [denoted in Appendix 1 as $\tau_0 + \tau_1$]

Letting (S, C, P) = (x, y, z) and setting the derivatives to all be identically equal to zero for equations (1) to (4), we find:

(2) & (5) $\Rightarrow$ $0 = \theta(1 - x - y - z) - \mu x - \lambda xy - \gamma z - \alpha x$ (6)

(3) $\Rightarrow$ $0 = \alpha x + \lambda xy - \tau((1 - x - y - z))y - \varphi_1 y + \varphi_2(1 - \pi)z$ (7)

(4) $\Rightarrow$ $0 = \varphi_1 y - \varphi_2 z$ (8)

Substituting in τ into equations (6) to (8) and rearranging we get:

$$0 = (\theta + \mu + \alpha)x + \theta y + (\theta + \gamma)z + \lambda xy - \theta \quad (9)$$

$$0 = \alpha x + (-A - B(1 - x - y - z))y + \lambda xy - \varphi_1 y + \varphi_2(1 - \pi)z \quad (10)$$

$$z = \frac{\varphi_1}{\varphi_2} y \quad (11)$$

This set of equations in this state could not could be solved explicitly straightaway, one further small simplification being required. In equation (10) we drop the y and z in (1 – x – y – z). The resultant change in (-A – B(1 – x – y – z)) is around three per cent, and the total change to the coefficient of y in this equation will be less than one per cent. This slight change has a negligible effect on the equilibrium solution, but makes the equations tractable. Equation (10) now becomes:

$$0 = \alpha x + (-A - B - \varphi_1)y + \varphi_2(1-\pi)z + (\lambda + B)xy \tag{10}$$

Substituting (11) into (10) we get:

$$0 = \alpha x + (-A - B - \varphi_1)y + \varphi_1(1-\pi)y + (\lambda + B)xy$$

$$\Rightarrow \qquad (\alpha + (\lambda + B)y)x = (A + B + \pi\varphi_1)y$$

$$\Rightarrow \qquad x = \frac{(A + B + \pi\varphi_1)y}{\alpha + (\lambda + B)y} \tag{12}$$

We now make some standard changes of variables:

$$k_1 = \theta + \mu + \alpha$$
$$k_2 = \theta$$
$$k_3 = \theta + \gamma$$
$$k_4 = \lambda$$
$$k_5 = A + B + \pi\varphi_1$$
$$k_6 = \alpha$$
$$k_7 = \lambda + B$$
$$k_8 = \frac{\varphi_1}{\varphi_2}$$

so that (9) and (12) respectively become:

$$0 = k_1 x + k_2 y + k_3 z + k_4 xy - k_2 \tag{9}$$

$$x = \frac{k_5 y}{k_6 + k_7 y} \tag{12}$$

Substituting (12) into (9) we get:

$$0 = \frac{k_5 y}{k_6 + k_7 y} k_1 + k_2 y + k_3 k_8 y + \frac{k_4 k_5 y}{k_6 + k_7 y} y - k_2$$

$$\Rightarrow \quad 0 = k_1 k_5 y + k_2 k_6 y + k_2 k_7 y^2 + k_3 k_8 k_6 y + k_3 k_8 k_7 y^2 + k_4 k_5 y^2 - k_2 k_6 - k_2 k_7$$

$$\Rightarrow \quad 0 = y^2\left[k_2k_7 + k_3k_8k_7 + k_4k_5\right] + y\left[k_1k_5 + k_2k_6 + k_3k_8k_6 - k_2k_7\right] - k_2k_6$$

This is a quadratic and so can be solved for y:

$$A' = k_2k_7 + k_3k_8k_7 + k_4k_5$$
$$B' = k_1k_5 + k_2k_6 + k_3k_8k_6 - k_2k_7$$
$$C' = -k_2k_6$$

$$y = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$
$$x = \frac{k_5 y}{k_6 + k_7 y}$$
$$z = k_8 y$$

These solutions were checked against numerical solutions that were derived with a fourth order Runge-Kutta algorithm, for many different sets of parameters. The two different sets of solutions were found to be in agreement in all cases, making us confident of their validity.

Examining the stability of the equilibrium solution

To examine the stability of the model as it is set up, we need to look at the community matrix, the set of all partial derivatives of the three defining equations above:

$$\underline{\underline{A}} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{pmatrix}$$

where

$$f = (\theta + \mu + \alpha)x + \theta y + (\theta + \gamma)z + \lambda xy + \theta$$
$$g = \alpha x + (-A - B - \varphi_1)y + \varphi_2(1 - \pi)z + (\lambda + B)xy$$
$$h = \varphi_1 y - \varphi_2 z$$

We then need to look at the eigenvalues of $\underline{\underline{A}}$ at the equilibrium solution (xs, ys, zs). This leads us to a cubic, $\mu^3 + a\mu^2 + b\mu + c = 0$ and for a stable equilibrium we need $\mathrm{Re}(\mu) < 0$. This is true if and only if the three Routh-Hurwitz conditions hold:

(i) $a > 0$
(ii) $c > 0$
(iii) $ab - c > 0$

Doing the calculations above, we find a, b and c:

$$a = R_1 - \lambda(x_s - y_s) + \partial_a$$
$$b = R_2 - \varphi_2\lambda(x_s - y_s) - R_3 x_s + R_4 y_s + \partial_b$$
$$c = R_5 - R_6 x_s + R_7 y_s + \partial_c$$

where

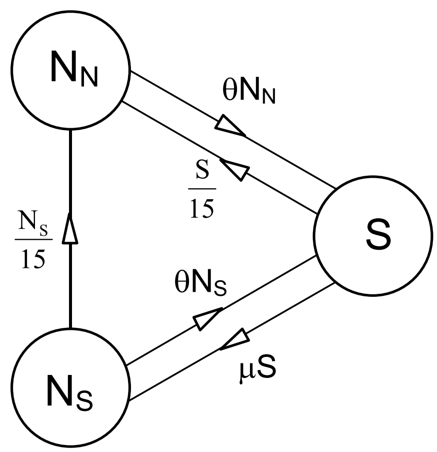
$$R_1 = \theta + \alpha + \mu + A + \varphi_1 + \varphi_2$$
$$R_2 = (\theta + \alpha + \mu)(A + \varphi_1) + \varphi_2(A + \varphi_1) + \alpha\theta + \varphi_2(\theta + \alpha + \mu) - \varphi_1\varphi_2(1 - \pi)$$
$$R_3 = \lambda(\theta + \mu)$$
$$R_4 = \lambda(\theta + A + \varphi_1)$$
$$R_5 = (\theta + \gamma)\alpha\varphi_1 + (\theta + \alpha + \mu)(A + \varphi_1)\varphi_2 + \theta\alpha\varphi_2 - (\theta + \alpha + \mu)\varphi_1\varphi_2(1 - \pi)$$
$$R_6 = \varphi_2\lambda(\theta + \mu)$$
$$R_7 = \lambda\left[(\theta + \alpha)\varphi_1 + \varphi_2(A + \varphi_1) + \theta\varphi_2 - \varphi_1\varphi_2(1 - \pi)\right]$$

$$\partial_a = -B(1 - x_s)$$
$$\partial_b = -B\left[\theta y_s + \lambda x_s y_s + \varphi_2(1 - x_s) + (\theta + \alpha + \mu)(1 - x_s) + \lambda y_s(1 - x_s)\right]$$
$$\partial_c = -B\left[\varphi_1(\theta + \gamma)y_s + (\theta + \alpha + \mu)\varphi_2(1 - x_s) + \varphi_2\lambda y_s(1 - x_s) + \theta\varphi_2 y_s + \varphi_2\lambda x_s y_s\right]$$

Using this analysis with the violent and property crime parameter sets calibrated with the late 1990s sets of parameter values, we find that the first and third Routh-Hurwitz conditions hold. However, the second condition does not, with $c < 0$. The value of this coefficient is only slightly less than zero, varying between 0 and –0.02. In practice, this does not appear to affect the stability of the system, as numerical simulation of the system shows. For a given set of parameters, wide variations can be made in the initial values of N, S, C and P, and the system converges to exactly identical solutions in each case.

Appendix 3. Estimating θ and μ

Making the assumption that people in the S category commit one crime per year, or more precisely that you enter S from N by committing a crime, then we can estimate θ by looking at the probability of never committing a crime over the 15 year period that we are looking at. The diagram below shows how we go about evaluating this:



In this sub-model we are focusing in on the N – S loop. N has been split into two categories so that now the interpretations are as follows:

N_N Not currently susceptible and never have been
N_S Not currently susceptible, but have been at some stage
S Susceptible

Explanation of the two flows leading into N_N:

(1) We would like the number of people in this system to remain constant.
(2) When a male in our target population reaches the age of 15 he enters N_N.
(3) Our age range for young males goes from 15 to 30, so each year 1/15th of the system's people leave.

(1), (2), and (3) imply that the two flows leading into N_N, must be as shown.

In order to solve this system, and find the proportions in the three categories as a function of θ and μ, we set up a system of linear equations where the system is assumed to be in equilibrium:

$$x = (1 - \theta)x + (1/15)(y+x)$$
$$y = (1 - (1/15) - \theta)y + \mu z$$
$$z = (x + y)\theta + (1 - (1/15) - \mu)z$$

where x, y and z equal N_N, N_S and S respectively.

These equations cannot be solved directly. Firstly, we must fix a ratio between θ and μ. Based on previous estimates of θ and μ, we take $\mu = 3.75\theta$. With this ratio fixed, the equilibrium solution is simply evaluated.

Now we can find values of θ for which we gain reasonable values of N_N, the proportion of people who never commit any crimes. Taking $\theta = 0.08$, leads to a value of $N_N = 0.45$, i.e. 45 per cent of people in our population are never susceptible to committing crime. These values seem reasonable.

Appendix 4. Implications of the crime rate in C on φ_1, the size of C and the number of crimes from C

φ_1 calculation: φ_1 = probability of individual in C being sent to prison (per year)
= 1 – probability of not being sent to prison
= 1 – (probability of not being sent to prison for 1 crime)crime rate
= $1-(1-\kappa)^n$ where n is the crime rate (crimes per year)

Here, κ is the probability of being sent to prison having committed one crime. This is calculated by taking the average number of recorded VATP crimes per year dividing by the average number of people sent to prison per year.

With φ_1 we can go on to calculate the implied numbers in C. We can do this as the ratio between φ_2 and φ_1 is (as an estimate) equal to the ratio between C and P, i.e.

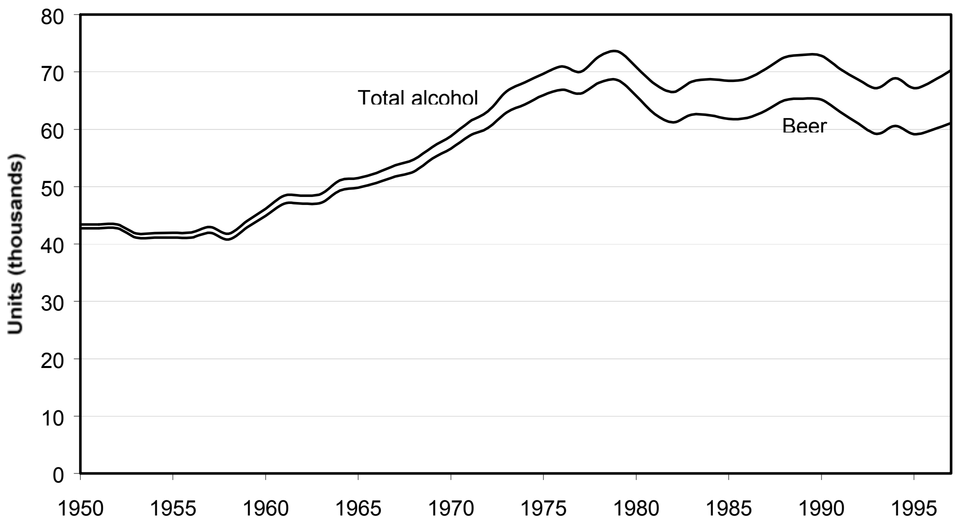
$$C \approx P \times \frac{\varphi_2}{\varphi_1}$$

Given the implied numbers in C and rate of crime for C, we also have the implied number of crimes for the men in C. The table below gives an idea of the implications of changing the rate crime in C for the late 90s value of κ.

Implications of crime rates for late 90s: $\kappa = \frac{1}{32.3} = 0.031$

Crime rate in C	Implied phi(1)	Implied C %	Implied number of crimes
0.5	0.016	27.9	209
0.6	0.019	23.3	209
0.7	0.022	20.0	210
0.8	0.025	17.5	210
0.9	0.028	15.6	210
1.0	0.031	14.0	211
1.5	0.046	9.4	212
2.0	0.061	7.1	214
2.5	0.076	5.8	216
3.0	0.090	4.8	217
3.5	0.104	4.2	219
4.0	0.118	3.7	221
4.5	0.132	3.3	222
5.0	0.146	3.0	224

Appendix 5. Trends in beer and total alcohol consumption



References

Austin, J.A. and Cohen, R.L. (1995) *Why Are Crime Rates Declining?* Briefing Report. National Council on Crime and Delinquency.

Becker, G. (1968) *Crime and Punishment: An Economic Approach.* Journal of Political Economy, vol.76, no.2, pp.443-78.

Currie, E. (1985) *Confronting Crime: An American Challenge*. New York: Pantheon Books.

Deadman, D. (2000) *Forecasting Trends in Recorded Crime*, research study for Home Office.

DiIulio Jnr., J.J. (1996) *Help Wanted: Economists, Crime and Public Policy*, Journal of Economic Perspectives, vol.10, no.1, pp.3-24.

Ehrlich, I. (1996) *Crime, Punishment and the Market for Offences*, Journal of Economic Perspectives, vol.10, no.1, pp.43-67.

Flood-Page, C. *et al* (2000) *Youth crime: Findings from the 1998/99 Youth Lifestyles Survey.* Home Office Research Study 209. London: Home Office.

Graham, J. and Bowling, B. (1995) *Young People and Crime*, Home Office Research Study 145. London: Home Office. Database at the Data Archive, University of Essex, http://dawww.essex.ac.uk, archive number 3812.

Marris, R.L. (2000) *Survey of the Research Literature on the Criminological and Economic Factors Influencing Crime Trends*. research study for the Home Office.

Murray, J.D. (1990) *Mathematical Biology*, Springer-Verlag.

Ormerod, P. and Campbell, M. (1997) *Social Interaction and the Dynamics of Crime*, Volterra research paper.

Ormerod, P. and Smith, L. (2000) *Random Matrix Theory and the Predictability of Time-Series Crime Data in the UK* research study for the Home Office.