The Safer Cities
Programme Impact Evaluation:

Problems and Progress[1]

BY PAUL EKBLOM

ABSTRACT

The Safer Cities Programme Impact Evaluation: Problems and Progress by Paul Ekblom. The Safer Cities Programme seeks to reduce crime and fear of crime, and to foster social and economic well-being, in 20 English cities and boroughs chosen for their level of social problems. The Programme, launched in 1988 by the UK Home Office as part of wider government action on cities, relies on local coordinators and modest funding of local crime prevention schemes using a broad range of preventive methods. The Research and Statistics Department of the Home Office is conducting an evaluation of the Programme's impact. This paper describes challenging methodological and logistical problems facing the evaluators and progress to date in meeting them. Research methods include analysis of police recorded crime statistics and crime survey data, both disaggregated to the small-area level. It is intended to measure the impact of large numbers of small schemes simultaneously to circumvent some well-known difficulties of evaluating single schemes. *(Studies on Crime and Crime Prevention Vol. 1. No. 1 1992. National Council for Crime Prevention).*

Keywords: crime prevention, evaluation, programme evaluation, population census, crime surveys, Geographic Information Systems (GIS), small-area studies, displacement.

*T*he launch, in March 1988, of the Home Office's Safer Cities Programme in England was the culmination of two processes: one in the realm of research and development, the other in the realm of politics and policy. The previous decade had seen a steady growth in research and development work in the field of crime prevention – notably involving the evolution of the situational approach to prevention, the creation in 1983 of the Home Office Crime Prevention Unit and the conduct of a body of action research recorded especially in the *Crime Prevention Unit Papers* series. Among the – now familiar – themes which emerged during the course of that work were the need to devise preventive strategies that dealt with specific, local crime problems, and the need, in identifying crime problems and devising and implementing solutions, to work collaboratively across the division of labour in what has become known as the 'multi-agency approach'. In 1984 a circular signed by the chief officials of five major government departments was sent out to a range of local organisations (including police, proba-

---

tion and local government bodies), commending to them the local, multi-agency approach to crime prevention. In 1986 the Home Office established the Five Towns initiative, in which a crime prevention coordinator was funded in each of five districts, with the aim of putting the local, multi-agency approach to the test. A retrospective evaluation has just been completed (Liddle and Bottoms, 1991).

On the broader front of politics and policy, a 'window of opportunity' (Tilley, 1991) appeared through which this decade of research and development work gained wider application. On winning the 1987 General Election the then Prime Minister declared an interest in taking action on the problems of the inner cities. Out of that aspiration was born the government's Action for Cities (AfC) Programme. It involved coordinated effort by a number of government departments including Environment, Trade and Industry and others, with the aim of improving social conditions in the inner cities primarily by regenerating the local economy. Much of the initiative involved coordinating existing effort and drew on existing funds. The Home Office's contribution, which involved new money, was the Safer Cities Programme (SCP).

The SCP is being implemented from the centre by the Home Office Crime Prevention Unit (CPU), which drew on existing experience in crime prevention activity, including the Five Towns initiative. Its objectives are to:

☐ reduce crime;
☐ lessen the fear of crime;
☐ create safer cities within which economic enterprise and community life can flourish.

Safer Cities is locally-based. Projects have been established in 20 local areas in England; so far 16 are running and the final 4 areas have just been selected. Each is located in an urban district area (a city or borough) which qualifies for funding through the Department of the Environment's Urban Programme (one of the major, big-spending AfC programmes which focuses on 57 cities and boroughs suffering from multiple social and economic problems). In each area, the project is guided by a local steering committee representing local government, police, probation, voluntary bodies and commerce, which sets priorities for the project and oversees the implementation of measures directed against crime. Each project is directed by a local coordinator with an assistant and a secretary whose salaries and running expenses are met by the Home Office. The Home Office also provides £250,000 per year per project for funding local organisations to implement preventive action (other funds may be levered in from the Urban Programme, for example). The kind of preventive action being funded covers an exceptionally wide spectrum. It includes, at local levels, for example, target-hardening, educational schemes, the installation of a single street light outside a retirement home, the upgrading of security on a large residential housing estate. Other action includes the targeting of vulnerable groups such as ethnic minorities, women or the elderly, a city-centre safety strategy and a city-wide publicity scheme. The scope of SCP action can be obtained from the progress reports published annually (e.g. Home Office, 1990).

It may be helpful briefly to mention some SCP terminology at this point. Each funded element of preventive action is known as a *scheme;* the structure in a given city or borough – the coordinator's team, the steering committee – and all its activities are known as a *project;* the SC *Programme* comprises in all 20 projects. The city or borough is called the *project area* to distinguish them from the *small areas* to be used in the evaluation design. Those small areas in which schemes occur are known as *action areas*.

In recent years the British government has paid increasing attention to issues of management and value-for-money in respect of its own activities, and, in conse-

quence, programme evaluation has begun to take its place in management practice in government departments. The Research and Statistics Department of the Home Office (RSD) is responsible for conducting an evaluation of the impact of the Safer Cities Programme on its objectives as just described. Other arrangements are being made for evaluating aspects of the Safer Cities *process* and *structure*.

The impact of the entire Programme, the 20 individual projects and the individual schemes (of which to date some 1600 have been started) are all legitimate subjects of interest. The RSD evaluation will devote most effort to the performance of the programme as a whole, both as a matter of principle and because of serious practical, methodological and resource constraints, mentioned at various points below. In addition, however, the Crime Prevention Unit (CPU), as managers of the Programme, are endeavouring to foster good-quality self-evaluation of schemes by coordinators. In order to provide background information for these evaluations, the coordinators should be able to draw on the data assembled by the RSD evaluation team.

The evaluation is generating a number of reports and other outputs to a time-scale which reflects the perennially-conflicting requirements of reliable results versus quick results. It is planned to produce the main and most conclusive findings in late 1992 / early 1993 but the RSD has been supplying 'quicker and dirtier' assessments before then to guide management of the Programme in the shorter term.

DATA

The evaluation will use three principal types of data, covering the nature and location of the preventive action, and indicators of its impact. *Action data* is being collected to measure what activity has been initiated at what place, against what offence types, by what method of prevention and – importantly – to what extent it has been success-

fully implemented in practice (for example, how many young people attended the club, how many houses were target-hardened). For reasons of economy and in order to avoid imposing extra demands on the implementors, this depends heavily on the SC Management Information System set up by CPU.

Outcome data is being collected to measure changes in crime and fear levels, impact on people's behaviour and their awareness of preventive activity. It comprises *police recorded crime statistics* and *surveys of residents*. Other aspects of data and data collection, including confounds and demographic covariate data, are mentioned subsequently.

DESIGN OF THE EVALUATION

The conventional evaluation of crime preventive action involves assessment of a single scheme, carefully selected and located in advance and subject to reasonably tight experimental control. But even under these ideal conditions, no evaluation can provide results which are free from uncertainty or beyond dispute (Ekblom, 1990). The RSD impact evaluation faces extremely difficult conditions which centre on having to detect the effects of a set of preventive schemes which are very diverse in size and nature (some being extremely modest), which may or may not be successfully implemented, which start up at different times, which possibly overlap, and whose locations are not merely scattered but unknown in advance. The effects have to be detected against a background of non-SCP preventive activity in the SC areas themselves and in the country as a whole together with a range of other AfC programmes which *in toto* dwarf the SCP in funding terms.

These harsh conditions reflect the real world of programme evaluation and make for a significant risk of *measurement failure* (Rosenbaum, 1986) – in other words the possibility that the SCP will have an impact but that the evaluation exercise will be unable to detect it. The basic design has been

developed to face up to this risk and to cope with it in as economical a way as possible – in particular to be extremely flexible, to adapt to the evolving nature of the SCP, to the unpredictable occurrence of confounds and to emergent problems with individual data sources. But it has been made quite clear to those with an interest in the results of the evaluation that the final results will inevitably be surrounded by a considerable margin of uncertainty (Ekblom, 1990).

THE BASIC DESIGN

The aim of the design is to detect change in crime, fear etc. which is associated with SC action in time and place, and not associated with other, confounding influences such as economic change or other AfC work. Police recorded crime statistics and survey data are being collected over a period which covers before and after the start of the preventive action; and the data in the SC action areas are to be compared against measures of background changes which might be mistaken for a 'Safer Cities effect'. In conventional research design terminology, the evaluation is, of course, quasi-experimental and has elements of the 'interrupted time series' design (with action and comparison series by month – mainly for the police recorded crime data) and the 'non-equivalent control group' design (i.e. before/after × action areas/comparison areas – mainly for the survey data) (Judd and Kenny, 1981).

Two complementary approaches to comparison will be used – each with different sources of error and together providing some protection against failure of one or other to deliver results. *External Comparison* will involve gathering crime and survey data from a set of cities and boroughs *not* participating in the SCP, but *like* the SCP areas themselves being members of the 57 Urban Programme Areas. The advantage of having a set of 8 or 10 cities/boroughs rather than relying on just one or two areas is that this tends to neutralise the effects of 'history'

(Campbell and Stanley, 1963) in single locations.

Internal Comparison, as the name implies, will look *within* SC project areas and compare before/after changes in places which have had SC action with those that have had none. Further details of how this will work are described in the next section.

PROBLEMS OF DETECTING CHANGES – THE LEVEL OF ANALYSIS

As said above, the focus of the evaluation is the impact of the SC Programme as a whole, but there was some considerable debate on the level of analysis most appropriate for this task. The two conventional alternatives we considered were the top-down approach and bottom-up. Top-down (or macro-level) would involve, for example, looking for changes in crime in whole project cities/boroughs; bottom-up (or micro-level) would involve evaluating a set of individual schemes and then drawing conclusions from these about the performance of the entire Programme. Neither approach was considered satisfactory.

Top-down. This approach would come up against the 'dilution' problem. Given that the modest sum of £ 250,000 per year is to be spent in cities up to the size of Birmingham (population nearly one million), the effect of preventive action, when viewed from the city perspective over the time-scale of this evaluation (up to early 1993), is likely to be very small in relation to the total baseline crime level and largely masked by random background fluctuation.

Bottom-up. For the major part of SC activity, individual schemes may be extremely small and/or diffused over a fairly wide area, so that detection of preventive effects is severely hampered by random fluctuation in crime. Evaluation of even quite large schemes may suffer from this problem. An analysis of year-to-year changes in street crime across each police beat in a London project area, for example, shows a very large

Figure 1:

Cross-sectional variation in change at the small area level - street crime, London

% change +/-

% change
Jan - June
1989/Jan -
June 1990

Police beat number
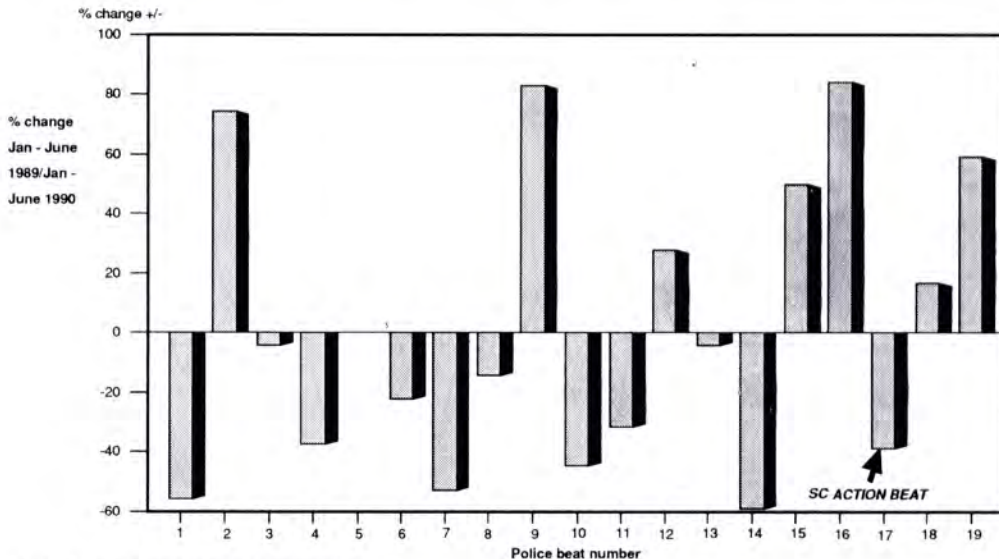
Source: Metropolitan police recorded allegations of crime

Note: The graph represents, for each police beat in one London Safer Cities project area,
the percent change in crime from the period January to June 1989 to the same
period in 1990.

spread of rises and falls (Figure 1). Within this area, the impact of one large scheme in beat no. 17 – which looks good in isolation – completely disappears. A time series of monthly recorded crime levels in one police beat in the West Midlands shows fluctuation of similar strength (Figure 2), which again throws doubt on the existence of a link between the implementation of a preventive scheme and the fall in crime that followed it. Despite these difficulties it should be possible to detect the impact of some individual schemes if they are large enough and sufficiently focused on specific areas. However, sole reliance on assessments of such schemes would, of course, severely bias the Programme evaluation since they only comprise a part of the whole effort – we would merely have an evaluation of 'Safer housing estates' rather than Safer Cities.

To cope with these problems a method called 'retrospective concentration' has been devised to give the greatest possible chance of detecting change. This involves, *in advance of preventive action*, collecting data over wide territories but disaggregated into small areas (such as police beats for the recorded crime data or Enumeration Districts from the population Census for the surveys); and then *retrospectively* homing in on the data from the specific small areas where action by then has been initiated. (As a useful 'by-product' of the retrospective concentration process, we will draw our Internal Comparison areas from the small areas that are left over.) The small 'action' areas identified can then be considered simultaneously in the search for impact, allowing the random 'noise' in the individual areas to cancel out, to leave the weak 'signal' of change systematically associated with preventive action to get through.

In effect, this approach combines features of top-down and bottom-up; it is a micro-analysis conducted over a macro-area. Making quite large numbers of small areas the units of analysis will provide a rare chance to get beyond the constraints and

Figure 2:

Longitudinal variation in change at the small area level

Total recorded crimes, one police beat, West Midlands
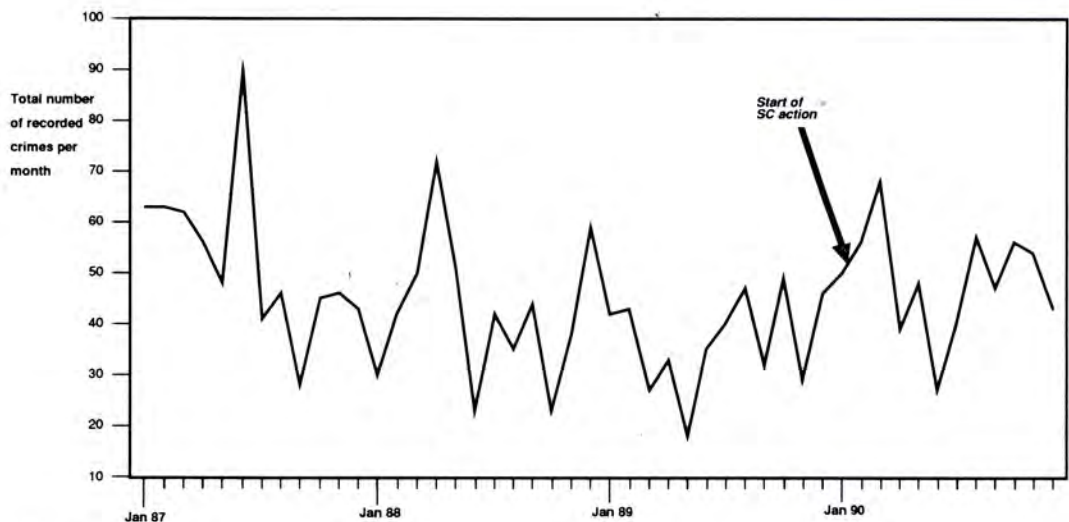
Source: West Midlands Police

risks of implementation failure and measurement failure associated with the traditional crime prevention evaluation which typically involves just a very small number of action and comparison areas. (Only a few such studies have been conducted in the crime/crime prevention area looking at change over time – e.g. Kohfeld et al. 1981; Kohfeld and Sprague, 1988; McDonald, 1986.) It can overcome a difficult paradox in evaluating crime prevention (cf. Ekblom, 1990) in which smaller schemes may be better implemented and have greater impact, but pose the greatest difficulties for the detection of any associated change. To take an example, Bennett (1990) had problems with implementation failure in his neighbourhood watch evaluation (the neighbourhood watch schemes he evaluated were chosen for their size in order to improve the chances of measuring change, but because of this self-same size they were not very thoroughly implemented on the ground). (He also had difficulty with measurement failure – his

single control area unluckily showed a fall in crime – and he now advocates a similar approach to that put forward here). Altogether, as far as detecting change goes, at least, the SCP evaluation becomes less of a risk and more of an opportunity.

PROBLEMS OF CONTROL – PROGRAMME BOUNDARIES AND CONFOUNDING INFLUENCES

SCP was not initiated in a vacuum: over recent years a considerable amount of local preventive activity has started to make its mark in most urban areas in England. Some of this activity is purely locally-funded and carried out; some comes under national AfC programmes such as the Urban Programme and Estates Action (refurbishment of housing estates, including some work on physical security of dwellings). This state of affairs leads to two linked problems – deciding what counts as SC action, and the more familiar question of confounds.

DECIDING WHAT COUNTS AS SC ACTION

It is clear enough that the impact of a particular scheme can be attributed to SCP if that scheme was wholly funded by SCP. But where preventive action was, say, merely given some limited assistance by the local coordinator through provision of advice or contacts, the extent to which SCP can claim credit for the action is rather less. Nevertheless, the SC contribution in these circumstances may not be negligible: part of the coordinator's job is to augment the direct SC funding by securing money or resources from elsewhere. Undoubtedly 'pure coordination' is a key task of the coordinator and the question here is whether the preventive action would have happened anyway in the absence of a coordinator. This problem is being tackled by paying careful attention to the definition of 'SC-owned' schemes in the Management Information System. A number of alternative categories of ownership are defined ranging from 'scheme funding and initiation by SC' to mere 'provision of advice or influence by SC'. Alternative analyses could be conducted on a) just the core schemes fully-funded by SCP, and b) the core plus the 'penumbra' schemes where SC can take credit for influencing or stimulating the schemes, but not for their funding.

COPING WITH CONFOUNDING EVENTS AND ACTIVITIES

Confounds in SC project areas. Bottom-up evaluation of SCP through assessment of individual schemes was rejected because of problems in detecting change due to smallness of areas, but assessment of individual schemes is also severely threatened by confounds. Urban areas, where SC action is targeted, are the focus of many other local and central government initiatives (including AfC), so that it is not surprising that many SC schemes are being located closely in time and space to other significant events and activities. These might include non-SC crime prevention schemes or housing improvement activities, which are likely to have their own impact on crime. Therefore, having succeeded in detecting reliable falls in crime associated with a particular SC scheme, it is almost inevitably the case that other activities can be found to compete for the credit. It is also often the case that it is impossible (even with detailed local data) to share out the credit quantitatively between the SC scheme and its 'competitors' (indeed it is possible that the SC scheme needed the prior presence of, say, a housing management scheme, to enable it to have any impact at all, in a kind of interaction effect). It is important to note that in view of the far greater funds available through the other AfC programmes, the *side effects* on crime of improvements in local housing or the local economy may be far greater than the *main effects* of the financially much more modest SCP, acting directly on offending.

The simultaneous consideration of large numbers of action areas may partially circumvent this problem if there is sufficient differential location of SC schemes and confound schemes. In other words, if, at the end of the day, we can find some small areas with SC action only, some with confound action only, some with both and some with neither, and all randomly-distributed within the same socio-demographic range, then we should be able to distinguish a SC effect at Programme level. In such circumstances confound action would merely act as one more source of background noise. It is, unfortunately, quite possible that confound action such as non-SC preventive schemes might somehow be systematically directed *away* from SCP action areas, for example by implementors such as the police or local government wanting to avoid overlap. This would confound the internal comparison process, bias the evaluation and lead to underestimation of the impact of SCP. Alternatively, non-SC preventive action and other AfC action might be differentially directed *towards* SC action areas because it is these areas which have serious multiple problems.

In all quasi-experimental designs it is important in interpreting results of the analysis to try to get a picture of the *assignment rules* for preventive action (Judd and Kenny, 1981). How were small areas within each project area selected for the placement of schemes? How in particular did the placement of SC-funded action relate to existing crime prevention effort in each project area and to other AfC work – was there no relation, was it differentially assigned to the same places, or to different ones? The rational, planned element of these processes is explicitly related to high rates of crime and other social problems, but it will be important to be alert to the intrusion of other systematic factors. This could be achieved by two approaches. First, by interviews with SC project coordinators and the central implementation team, plus organisations responsible for placing the confound action on the ground; and second, to seek as far as is practical to map the location and nature of the confound activity within the project areas to see whether such differential distribution was occurring, and to try to take account of it in analysis.

CONFOUNDS IN EXTERNAL COMPARISON

Equivalent problems exist with the External Comparison process. One of the clear assignment rules for selecting SC project areas within the 57 Urban Programme areas has been a preference for cities and boroughs with higher crime rates. Non-SC preventive action is certainly differentially finding its way to urban areas outside the SCP. For example, the organisation *Crime Concern* is setting up citywide strategic crime reduction projects rather similar to those of SCP in some of the 57 Urban Programme areas, and is obviously avoiding those already chosen for SCP. Cities and boroughs containing such strategic crime prevention projects can be weeded out of the comparison, preferably in prospect, as the next sections describe,

but if necessary in retrospect. Likewise, the distribution of other AfC action between SC cities/boroughs and the other Urban Programme areas will be examined and taken account of in analysis.

However successfully we cope with these confound problems (and however fortunate or unfortunate the geographical distribution of confounds happens to be) it has to be said that they place *fundamental limits* on the probability of the evaluation leading to a clearcut conclusion, as they do for evaluations of any individual component of the whole AfC package. In this context we are faced with a situation in which one evaluator's confound is another evaluator's programme. It is worth pointing out that the Home Office, along with other government departments involved, is contributing funds to an evaluation of the impact of AfC as a whole.

VARIATIONS ON A THEME: CRIME DATA AND SURVEY DATA

While we plan to follow broadly similar design principles with the recorded crime data and the surveys considerations of methodology, cost and logistics make design and analysis somewhat different.

Recorded crime data strategy

The advantages of recorded crime data are that they are collected as a matter of routine, over a long period, and readily-disaggregated into reasonably small areas – in particular, the police beat. The beat, the unit of uniformed foot patrol, usually covers a few thousand households. Because it is a practical unit of police operations rather than a unit of analysis it is far from ideal for small-area research (Ellis, 1990), being somewhat large, arbitrary in relation to crime patterns and community or administrative boundaries and liable to be changed at any time. In quite a few cases only paper

records have been available, leading to burdensome data-gathering. But there are, for the present, no alternatives for an evaluation which cannot afford the time and effort required to map each scheme precisely and identify recorded crime incidents from just those exact areas. In the medium term, though, all police forces will move towards address-referenced computerised logging of all individual crime incidents. This would enable the kind of data we require to be obtained and aggregated into appropriate standardised small areas very simply, quickly and cheaply, and might even make possible the move from standardised small areas to precisely-mapped action areas of the kind just mentioned.

The plan is to obtain crime data by beat (about 1,200 in all), by monthly time series (from January 1987 predating the start of SCP action on the ground by about 2 years), by crime type (12 major categories), for all SC project areas. Each beat will be associated with a residential population figure to produce per capita rates, and with other demographic information. Because beats bear no relationship to local government administrative territories, their population has to be worked out by use of a geographical information system which 'fills' the boundaries of each beat (represented on computer) by the relevant Census Enumeration Districts, the smallest geographical disaggregation for which population Census data is obtainable – usually 100–150 households.

The monthly beat data was wanted by the SC project coordinators in planning preventive activity and conducting their own assessments of schemes. It was the intention to exploit this same data for the purposes of evaluation, thereby reducing the burden on coordinators and evaluation team alike. For the External Comparison, there was unfortunately no ready-made *monthly* data source (only an annual series that would have been too crude for the kind of analysis planned), so we had to construct a monthly time series that covered comparable areas and behaved

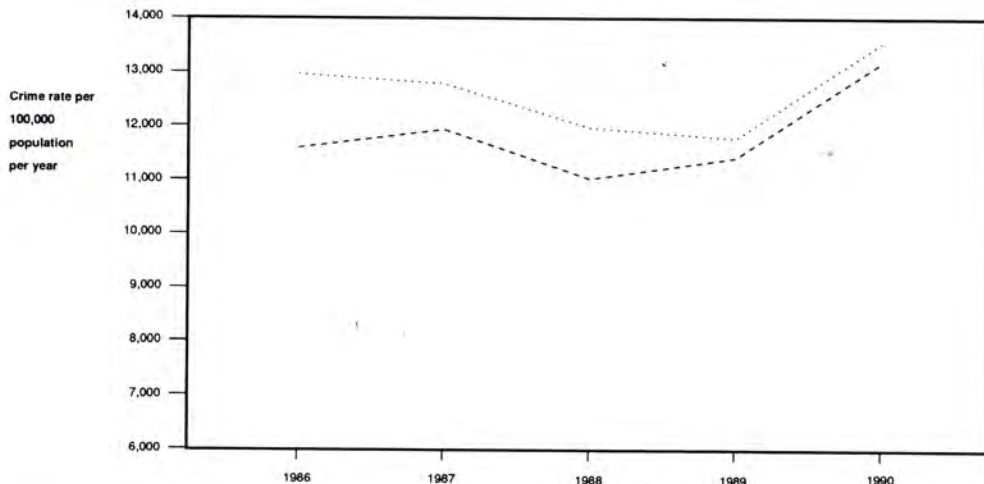in a fairly similar way to the SC project areas.

Starting with the 57 cities and boroughs in the Urban Programme, we eliminated those which had SC projects already in existence or planned, or if they had an obvious confound – in the form of major strategic preventive activity originating elsewhere. Next, we used Craig's (1985) classification of British cities and boroughs based on cluster analysis of 1981 Census data to stratify the 20 SC project areas into several *families* (including an amalgamated one for London boroughs). We then identified sets of likely comparison cities and boroughs, within the London family and each of the 3 provincial families separately.

The SCP's implementors had, by and large, selected the 20 project areas from the 57 Urban Programme areas by starting at the upper end of the scale of per capita recorded crime rates and working downwards, so the various non-SC cities and boroughs within each family tended to average out at lower crime rates than their SC equivalents. Therefore, we reduced these sets of possible comparison cities and boroughs by filtering out those at the lower end of the crime-rate scale. This was done on the basis of the crime rate for each city or borough averaged over a five year period, to minimise the risk of our capitalising on similarity caused merely by random fluctuation over a shorter term. The average of the External Comparison set was weighted to make its family composition reflect that of the SC set.

Ultimately, ten cities and boroughs were chosen for External Comparison – sufficient to provide some protection against the vagaries of 'history' in individual locations. Given that by the end of the selection process there were relatively few suitable cities and boroughs left from which to choose, the final ten External Comparison cities and boroughs and their 20 SC equivalents did not match perfectly in terms of crime rates, trend or variability (standard deviation) over time, but they were close enough to

Figure 3:

Recorded crime data - Safer Cities and External Comparison

Crime rate per
100,000
population
per year

(graph with y-axis from 6,000 to 14,000, x-axis years 1986 to 1990)

Source: Home Office

Note: Construction of External Comparison
series described in text

Key: ···· Weighted mean yearly rate for 10 External Comparison areas (overall mean = 12,609, sd = 653)
- - - Mean yearly rate for 20 Safer Cities project areas (overall mean = 11,833, sd = 742)

provide a reasonable basis for detecting differential change. Figure 3 shows the five-yearly plot for all 20 SC project areas combined, plus all 10 External Comparison areas combined. This shows that our matching averages were not achieved by trends that were crossing over. (It should be noted that we did not expect to see any fall in the overall crime rate of the SC set relative to the External Comparison set, as the overall SC figures have not yet been 'retrospectively concentrated' and besides, sufficient amounts of preventive action have not been in place for long enough for much impact to be expected.) Having selected the External Comparison areas, the next step will be to collect monthly crime data to provide a time series from January 1987 equivalent to the SC series. The use to which this time series will be put is described in the section on analysis.

SURVEY DATA AND DESIGN[2]

Crime surveys originated in attempts to derive reliable and accurate estimates of the prevalence and incidence of victimisation, whether at the national level (e.g. British Crime Survey) or the local level. Adapting the crime survey for the purpose of programme evaluation has been challenging, despite helpful earlier work (Rosenbaum, 1986; Hope and Dowds, 1987; Bennett, 1990; Hope and Foster, in preparation). The advantages of using surveys in evaluation include the ability to measure impact on variables such as fear and routine activities related to quality of life and the local economy (e.g. going out for leisure, shopping or work). A further advantage is the avoidance of a major worry about using police recorded crime statistics as the sole indicator of impact, namely that the reporting

---

2. A more detailed account of the design of the main survey is in Ekblom (1991).

rate by victims to the police may increase as a result of the preventive action, thus giving an underestimate of the true change. (There does, however, remain the possibility that experience of preventive action causes people to report more victimisations to crime survey interviewers and not just to the police).

Conventional crime surveys are expensive and deciding on the right number of interviews in advance in order to be sure of detecting changes in prevalence of victimisation is extremely difficult – numbers required are usually very large because of the relative rarity of criminal incidents among the surveyed population and there is ample room for getting it expensively wrong. In fact, our range of working assumptions entered into the 'sensitivity' formula meant that we had to choose within an eightfold range of sample size. The top of the range was impossibly large for our budget and even then there was a reasonable probability of it not being sufficient to avoid measurement failure. In view of this, we lowered our aspirations and opted for a survey in which much reliance is placed on topics for which change is easier to detect because they are not dependent on the occurrence of rare incidents. As examples mention can be made of respondents' perceptions of change in crime and safety and change in attitudes to prevention, in protective behaviour (such as locking windows), and in 'amenity behaviour' (going out for leisure, shopping). Changes in prevalence of crime, nuisance and racial harassment are still being measured, but the survey part of the evaluation will not be wasted if these fail to reach significance. In some respects, anyway, reductions in crime prevalence are not the final word in this evaluation. Given that SCP is serving the wider aim of facilitating AfC, it could be argued that the 'feel good, act confident' factors are more important in contributing to the improvement in social and economic life of a city than the objective crime level.

To save interview time, the victim form (usually employed to collect extra details on crimes suffered) was omitted, crime prevalence being measured using the standard victimization 'screening' questions. The same approach was used by Bennett (1990) in his evaluation of neighbourhood watch. This meant that several items usually located in the victim form had to be brought forward into what was normally the 'main' part of the questionnaire, including perceptions of racial motivation and reporting to the police. The latter information was judged important to obtain as a check on the occurrence of changes in reporting rates between the before and after surveys, or even to use in order to correct the recorded crime figures in the relevant after-period. Rather than relying solely on respondents who described victimization incidents to us, in order to get a more reliable (but arguably less valid) estimate of reporting changes, we included questions about reporting intentions concerning several *hypothetical* incidents of common offences. Of necessity these had to be quite detailed in their specification of the events to be considered.

The surveys are to involve an 'embedded-panel' design (i.e. the after-survey will involve interviews with as many respondents as possible who participated in the before-survey, plus a fresh set of randomly-selected respondents in the same small areas to make up numbers). This increases sensitivity to change and allows for efficient analysis of impact on both individuals and on areas (cf. Bennett, 1990).

THE SMALL AREA SAMPLING STRATEGY

It was decided to conduct surveys in 11 SC projects out of 16 operational at the time. Census Enumeration Districts (EDs), at about 150 households each, were to constitute the basic small area unit used in the survey, offering a much finer-grained analy-

sis than with the police beats. Within the 11 project areas, we had to choose the small areas in which to interview. The most difficult problem was not knowing, at the time of planning the survey, the location of future preventive schemes because, unlike with the police recorded crime data, it was obviously impossible to collect survey data for every small area in every project area. At the time of planning the survey, most SC coordinators could identify with confidence one or two distinct residential neighbourhoods where their efforts were going to be concentrated, so focusing on these was a possibility.

However, they were not by themselves enough i) to generate a sufficient number of interviews; ii) to give a sufficiently large number of schemes covered to allow for wastage of some through implementation failure; and iii) to be representative of the broad picture of SC schemes – avoiding the 'safer housing estates' problem mentioned already.

In order to maximize the chance of "hitting" schemes whilst at the same time obtaining a sample of respondents with a broad experience of SC action, we split the Safer Cities survey into two subsamples within each project area: the *SC Targeted* subsample and the *SC Citywide* subsample. The *SC Targeted* subsample focused on the neighbourhoods just mentioned – geographically compact areas (generally housing estates) in each of which the coordinators told us substantial action was very likely to occur, probably involving the implementation of several schemes together. In all, 19 targeted 'action areas' were identified across 10 of the 11 SC project areas surveyed.
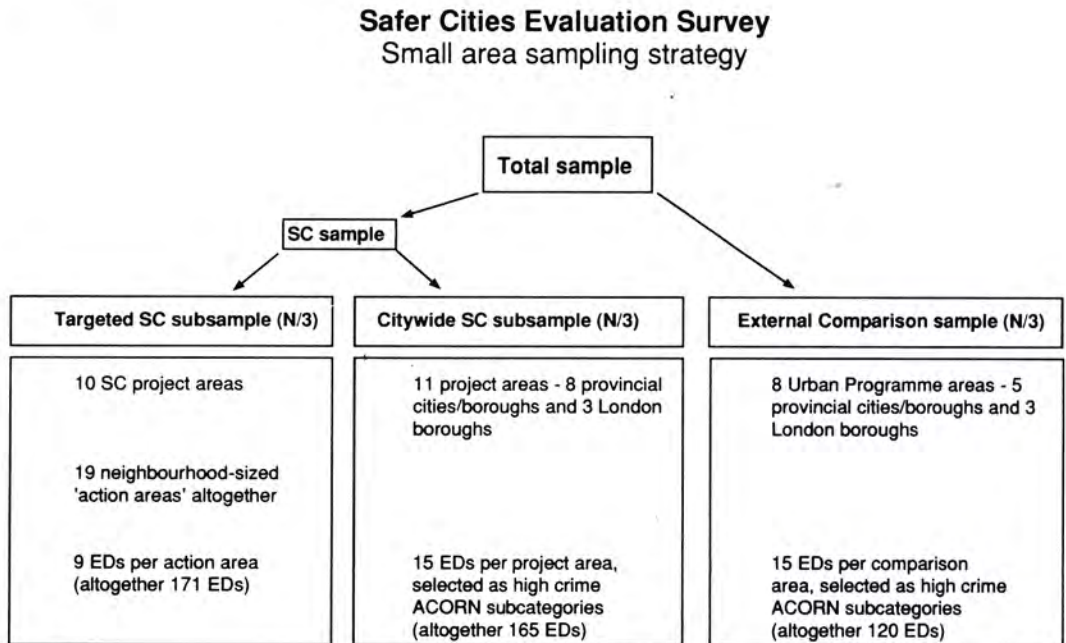
For the *SC Citywide* subsample the main aim was to choose a set of small areas each of which would be *the kind of area to which SC coordinators might consider directing preventive schemes*. In practice, this attempt to guess the coordinators' future assignment rules meant finding small areas with high per-capita crime rates. Across each project area we required a fairly wide scatter and a suffi-

cient number to maximize the chance of "hitting" well-implemented localised schemes, to provide a reasonably representative base for city centre questions and to give a good chance of drawing in vulnerable groups such as the elderly and ethnic minority residents.

To identify high-crime rate small areas we decided to use the ACORN typology of residential areas – based on a cluster analysis of Census variables – rather than recorded crime levels. ACORN is described in the 1988 British Crime Survey report (Mayhew et al. 1989), which confirmed the finding of the 1984 British Crime Survey (Hough and Mayhew, 1985) that ACORN categorisation can be used as a fairly good predictor of *relative* crime levels of small areas. Methodologically speaking, ACORN was the preferred proxy indicator because we wanted the survey to yield outcome measures which were as independent as possible of the other outcome measure – the police crime data. We also had a logistical problem to solve. Collecting beat-based crime data for the SC project areas had proved to be a very time-consuming exercise, as had linking up the beats to demographic information to obtain the population base. We had no wish to repeat this at a small area level for the External Comparison cities and boroughs, where we needed to sample small areas on a comparable basis to the SC Citywide subsample. Again, ACORN – being available commercially for all EDs throughout the country – served this purpose.

We started with the intention of focusing on the 5 out of 11 ACORN categories that the BCS identified as at highest risk of crime. However, on examining their distribution in the project areas we found, especially in the London borough projects, that they covered most of the boroughs and thus were no help in efficient targeting of the survey. We therefore turned to the 38 ACORN subcategories to be guided in on the highest-crime small areas. We conducted a special analysis in which, in order to provide sufficiently large sub-cells, we

Figure 4:

Safer Cities Evaluation Survey
Small area sampling strategy

Total sample

SC sample

Targeted SC subsample (N/3)	Citywide SC subsample (N/3)	External Comparison sample (N/3)
10 SC project areas	11 project areas - 8 provincial cities/boroughs and 3 London boroughs	8 Urban Programme areas - 5 provincial cities/boroughs and 3 London boroughs
19 neighbourhood-sized 'action areas' altogether		
9 EDs per action area (altogether 171 EDs)	15 EDs per project area, selected as high crime ACORN subcategories (altogether 165 EDs)	15 EDs per comparison area, selected as high crime ACORN subcategories (altogether 120 EDs)

combined data from the 1984 and 1988 British Crime Surveys to give a total N of over 21,000. All this enabled us to focus the Citywide sampling frame on 11 out of the 38 subcategories, equivalent to 39% of the total population in the 11 SC project areas in question, with the least focused of these – London Borough of Wandsworth – at 61% of the population.

EXTERNAL COMPARISON AREAS

Eight cities and boroughs were selected for External Comparison by the same process as for the recorded crime data, already described. Selecting small areas within the External Comparison cities and boroughs was relatively straightforward by design, since we used the same ACORN subcategories as for the SC citywide subsample.

Figure 4 shows the overall small area sampling strategy. In every case 22 addresses were identified for interview per ED,

of which it was hoped to achieve about 15 interviews. For the SC Citywide subsample, interviews were located in 15 EDs in each of the 11 SC project areas. Within each project area the relevant ACORN subcategories were represented roughly in proportion to their frequency of occurrence there. The same applied to External Comparison areas. For the SC Targeted subsample, 9 EDs were surveyed for each of the 19 targeted action areas.

For the main before-survey (conducted in late 1990) N was 4,388, divided roughly equally between the 3 subsamples of SC Citywide, SC Targeted and EC. Given such an ambitious design it is reassuring that the findings show that the three subsamples are rather close in terms of demographic measures, rates of crime prevalence and fear scores. This is all the more gratifying considering that the area sampling strategy was based on Census information that was nine years out of date. The response rate, at 58% overall, was however rather low – although

this could be expected to be a problem in the kind of inner-city areas in which we were interviewing.

PLANS FOR ANALYSIS

Plans for analysis are currently still evolving – the design was deliberately constructed to allow for a wide range of possibilities. One key factor which should become clear reasonably soon, now that the Management Information System is operational, is the geographical distribution of schemes within each project area – the extent to which they are scattered, clustered or overlapping, and the extent to which they are focused on small areas, are citywide or directed towards diffuse target groups such as the elderly.

The task of analysis will (at the risk of restating the function of the design) centre on detecting differential change (relative to Internal and External Comparison areas) in the small areas where SC action has been implemented, taking account of confounding factors and demographic covariation. Given that the evaluation is intended to guide decision-making by government regarding the continuance or expansion of SCP, we will not be employing an arbitrary and fixed 'p = .05' cutoff-point for significance. As far as possible we will attempt to quantify the SCP's impact on the various outcome measures, although the margins of error will be large. Statistical uncertainty surrounding the estimates of the size of impact can itself be quantified as confidence limits, but the non-statistical sources of uncertainty, relating for example to judgement as to the boundaries of SC activity, and to effects of confounds such as other AfC activity, will in some respects be unquantifiable (Ekblom, 1990).

The basic retrospective concentration process will be achieved by multivariate analysis in which SC action areas are identified by a dummy variable or – if there is substantial geographical overlap of schemes – by some kind of treatment score. The lat-ter would enable a 'dose-response' analysis to be conducted to assess whether more action in an area systematically leads to more impact or whether, perhaps, there is 'synergy' – greater impact than expected from the sum of individual schemes. Retrospective concentration could be further enhanced by identifying only those areas with schemes that have been judged as properly implemented in practice, and/or with schemes which the project coordinators judge to be *capable of having an impact* by the time of this evaluation (both are items which will be judged by coordinators for each scheme and routinely entered by them on the Management Information System) – educational initiatives, for example, may take considerably longer to have their effect. It will also be sensible to look at more specific input/outcome relationships, e.g. impact on burglary for only those areas with burglary schemes. This last kind of analysis will obviously be of interest in its own right, as well as supporting the general principle of looking for change only where change might reasonably be expected to be found.

DISPLACEMENT

With the Internal Comparison in particular, there will be a need to deal with the possible *confounding* effect of geographical displacement (Ekblom, 1990). Some small areas, whilst having no preventive action themselves, will be *adjacent* to those which do. They might be expected to show a *rise* in crime or fear if displacement had occurred into their territory; as such, this would mimic a rise in the common background rates, and lead us to overestimate any falls in the action areas. Displacement from non-SC preventive action might also have to be taken into account in identifying small areas likely to be influenced by confounds. Identification of small areas vulnerable to displacement will be aided by the use of a Geographic Information System (Arc/Info on a VAX 3100 Workstation) in which the

location of the preventive action will be represented topographically.

ANALYSIS OF RECORDED CRIME DATA

Analysis will probably involve calculating the level of crime *expected* in the action areas – on the basis of past trends in the action areas before the action was inaugurated and continuing trends in the Internal and External Comparison areas. This produces an estimate of what might have been had there been no SC action, and can be compared with the *actual* level of crime observed after the action began. If our police recorded crime data set is sufficient (i.e. at least 5 years' monthly data) we expect to employ multivariate time series techniques. If not, we may fall back on establishing before and after blocks of time for comparison through regression techniques.

ANALYSIS OF SURVEY DATA

The search for differential change will be done using as units of analysis the set of over 400 EDs where before and after interviews will have been conducted. We will use multiple regression both as the method of 'retrospective concentration' on where the preventive action occurred and to take account of confounds and demographic covariates statistically.

Regression would involve pooling the before and after surveys and probably using the technique of 'regression adjustment' in which the after-score is regressed on the before-score and covariates (Judd and Kenny, 1981). To improve the power of tests we will also pool the cities and boroughs, and pool the three main subsamples (Targeted, Citywide and External Comparison). We will use, or create, something like the following list of variables for each ED to incorporate in the regression equations:

☐ Outcome measures (fear, protective behaviour, victimization).

☐ Wave (before measure versus after measure).
☐ Treatment (preventive action implemented in the ED).
☐ Treatment dosage.
☐ Demographic covariates – for individual respondents and for EDs.
☐ Presence of confounds (non-SC crime prevention activity; other AfC activity).
☐ SC Targeted, SC Citywide or External Comparison.
☐ London borough versus provincial term.
☐ Displacement term.

The regression analysis will proceed in parallel from the panel perspective and the cross-sectional perspective. We might further explore the interface between individuals, small areas and the city/borough level through multi-level modeling. Unlike the British Crime Survey and local crime surveys whose aim is to yield a representative estimate of prevalence, we shall not apply weights – which would probably interfere with regression. Instead we will incorporate the factors on which weighting normally occurs (e.g. household size) into the regression process which will adjust statistically as necessary. An attempt will also be made to take account of the sampling error associated with each small area, something that seems rarely to be done.

CONCLUSIONS

The Safer Cities Programme evaluation is itself experimental. The requirement to evaluate the impact of SCP has led us into unknown territory: we have had to develop a rather different approach from that usually employed in the evaluation of individual preventive schemes. Simultaneous consideration of large numbers of preventive schemes at the small area level offers the chance of bypassing some difficult methodological problems and risks – but, obviously, introduces different problems of its own. The traditional single-scheme evaluation resembles a craft-based, labour-inten-

sive activity where researchers go out into the field and collect a great deal of detailed data about the one scheme and its context. The present evaluation more closely resembles an industrial process, with standardised, and coarser-grained, data being collected for well over a thousand schemes and their associated small areas.

In the past, evaluators were often able to 'parachute in' and assess an initiative on their own terms. Increasing government interest in evaluation has meant that evaluators and the designers and administrators of government policy are now being drawn together far more closely than before. The over-riding need to adapt the present evaluation to the function of informing major 'real-time' policy decisions about SCP has meant that a number of parameters have had to be changed from their 'standard academic setting'. In particular, we have attempted to strike a more even balance between the risk of Type I and Type II errors (mistakenly inferring impact versus mistakenly inferring no impact). The design goes as far as it can – without introducing bias – towards maximising the chance of finding a Safer Cities effect and minimising the probability of *mistakenly* concluding that 'nothing works' – for example by following the principle of looking for the Safer Cities effect only where and when we might expect to find it.

We have also attempted to be honest regarding the uncertainty that will inevitably surround the final results, and frank about the risks of measurement failure which we have striven to keep to a minimum within the resources available to us. In absolute terms, those resources are not small – about £ 600,000 plus staff expenses. But in relative terms, this represents only about 4% of the SCP implementation budget of £ 24m actual and planned spend up to March 1993 – the equivalent of one Safer City project. (The percentage is significantly smaller if one also takes into account the value of the SCP's coordination activity and its leverage of other funds and resources – about £ 7.5m to date.) It is difficult to establish a norm against which to judge this proportion – there having been little prior experience in crime prevention programme evaluations at least – and the appropriate norm would have to depend on the function of the evaluation, principally whether it merely served the purpose of retrospective accounting versus that of guiding significant policy decision-making in the future. The function of the SCP evaluation is clearly the latter – contributing to the development of SCP in subsequent years, in particular through the decision to continue with the Programme and the direction of that continuance. There should also be significant side-benefits, such as creating a large small-area database of crime and preventive action which could be used for example to investigate displacement, and developing a range of new techniques for evaluation in crime prevention including the use of geographic information systems.

REFERENCES

Bennett, T. (1990). *Evaluating Neighbourhood Watch*. Aldershot: Gower.

Campbell, D. & Stanley, J. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Craig, J. (1985). *A 1981 Socio-economic Classification of Local and Health Authorities of Great Britain*. OPCS Studies of Medical and Population Subjects 48. London: HMSO.

Ekblom, P. (1990). 'Evaluating crime prevention: the management of uncertainty'. In:

Kemp, C, ed. *Current Issues in Criminological Research vol. 2*. Bristol: Bristol Centre for Criminal Justice.

Ekblom, P. (1991). *High crime areas, surveys and evaluation: the Safer Cities Programme*. Paper presented at British Criminology Conference, York, 1991. (Available from the author.)

Ellis, T. (1990). *Police beat boundaries and crime: a GIS solution?* Paper presented at European Geographical Information Systems conference Amsterdam. (Available from the author.)

Home Office (1990). *Safer Cities Progress Report 1989–90*. London: Home Office.

Hope, T. & Dowds, E. A. (1987). 'The use of local surveys in evaluation research: examples from community crime prevention'. (Unpublished Home Office paper presented to British Criminology Conference 1987.)

Hope, T. & Foster, J. Home Office Research Study on Priority Estates Project's impact on crime. (In preparation.)

Hough, M. & Mayhew, P. (1985). *Taking Account of Crime: Findings from the Second British Crime Survey*. Home Office Research Study 85. London: HMSO.

Judd, C. & Kenny, D. (1981). *Estimating the Effects of Social Interventions*. Cambridge: Cambridge University Press.

Kohfeld, B., Salert, B. & Schoenberg, S. (1981). 'Neighbourhood associations and urban crime'. *Community Action Nov/Dec*: 37–43.

Kohfeld, B. & Sprague, J. (1988). *The relative invariance of predictive models for crime distributions across census tracts and census block groups*. Unpublished paper presented at annual conference of the American Society of Criminology, Chicago.

Liddle, M. & Bottoms, A. (1991). *Implementing circular 8/84: a retrospective assessment of the Five Towns initiative*. Unpublished report to Home Office.

McDonald, S. (1986). 'Does gentrification affect crime rates?'. In: Reiss, A & Tonry, M, eds. *Communities and Crime*. Chicago: University of Chicago Press.

Mayhew, P., Elliott, D. & Dowds, E. A. (1989). *The 1988 British Crime Survey*. Home Office Research Study 111. London. HMSO.

Rosenbaum, D. ed. (1986). *Community Crime Prevention: Does it Work?* Beverly Hills, California: Sage.

Tilley, N. (1991). *Opportunity knocks! – Crime prevention and the Safer Cities Story*. Unpublished paper presented at Social Policy Association annual conference, Nottingham University. (Available from the author at Nottingham Polytechnic.)

Paul Ekblom
Home Office Research and Planning Unit
50 Queen Anne's Gate
London SW1H 9AT
England