Crime surveys as tools of policy-making

Killias, M

Postprint available at:
http://www.zora.uzh.ch

Posted at the Zurich Open Repository and Archive, University of Zurich.
http://www.zora.uzh.ch

Originally published at:
Crime Surveys as Tools of Policy-making

Martin Killias
University of Zurich.
Chair of the “European Sourcebook of Crime and Criminal Justice Statistics” Project
International comparisons: an impossible adventure?  

We all agree that international comparative research is necessary. Policy-makers need benchmarks because it is important to know where we stand now in comparison to twenty years ago, or why we are worse or better off than others. Policy makers, at least if they are interested in the wider picture, also want to have these differences explained. Why are we where we are, why one approach does or does not work, what we currently do, where may we be in ten years if present trends continue, and what can we do to influence future changes. This has actually been a concern of policy-makers over long periods. In the area of crime, the first obvious need is to know the development of crime across space and time. Quételet began comparing justice and conviction data in the 1830’s. He made the assumption that the so-called “dark figure” would remain constant across time and space. However, this assumption is no longer shared by the scientific community, and has now been contested by Quételet’s contemporaries, such as the Genevan de Candolle who has since correctly identified the various sources of shortcomings in crime statistics.

There are obvious obstacles to legal definitions – these are the statistical problems with which all lawyers are familiar. However, when it comes to comparing police data, legal definitions are not that important. Much more problematic are the different ways in which police forces all over the world classify offences for statistical purposes. Even if offences are not uniformly defined across Europe, the question of whether the police count offences at the time they are reported, or after a successful prosecution (input vs. output statistics), is far more important. The situation is even worse when we look at how repeat, and particularly serial offences, are recorded. Sweden, for example, in cases of domestic violence or sexual abuse, counts all past incidences as having been committed at once, if necessary by extrapolation. If, for example, a woman reports having been assaulted by her spouse, the police will ask whether this incident is unique or whether similar events have happened before. In the latter case, they will ask how often and how long the relationship lasted – and finally record a possible 100 incidences of domestic assault if the vic-

1 References and data can be obtained from the author. Write to martin.killias@rwi.uzh.ch
tim complains having been assaulted once per week in a relationship that lasted for two years. All these difficulties have plagued statistical comparisons – particularly across international borders – since such endeavours began. This led fairly quickly to general pessimism. When we started the European Sourcebook of Crime and Criminal Justice Statistics, many observers felt it was pointless to undertake a compilation of statistics across Europe. For my part, I was more optimistic because I was familiar with the American Sourcebook of Criminal Justice Statistics. As America has fifty States each with their own legislation, I could not accept that we might fail where they have succeeded.

Although it may sound strange, I must tell you that after our first meeting we would probably have gone home never to meet again had our Committee’s secretary not arranged an excursion to a beautiful and majestic castle after which we visited a congenial inn where we had excellent Alsace specialties. After that evening, everyone was feeling more relaxed and the general conclusion was that, even if comparing statistics sounded impossible to many of our number, we should continue to meet at Strasbourg in the future. So, an excellent dinner and congenial company saved the project, and pessimism rapidly faded away once we delved more deeply into the matter. Indeed, we understood that comparing statistics implies comparing offence definitions and counting conventions. Countries using output statistics will obviously have lower crime rates than countries which count offences at the input stage, or countries which count serial offences (such as repeated domestic abuse or drug trafficking) as separate incidents. These will have dramatically higher rates than those who count all multiple crimes as one single event. Marcelo Aebi (University of Lausanne and Autonomous University of Barcelona) and Hanns von Hofer (University of Stockholm) have documented the effects of such conventions, based on data from the European Sourcebook of Crime and Criminal Justice Statistics (ESB). For example, according to the International Crime Victimisation Surveys, Sweden has, according to the ESB about seven times more police-recorded cases of rape per 100,000 of the population than Spain, whereas the two countries have quite similar rates of rape and sexual assault.

Crime surveys and the renaissance of international comparative research

This brings me to crime victimisation surveys. These studies began in the 1960’s when President Lyndon B. Johnson faced a difficult re-election campaign. His republican opponent, Arizona Senator Barry Goldwater, had made crime an issue, probably for the first time in an electoral campaign in a western country. President Johnson, felt rather unprepared on that issue, as are liberals and left wing politicians today when confronted with the problem of crime. He did what politicians always in these circumstances - he appointed a committee. This committee was chaired by his justice minister, Attorney-General Katzenbach, who published a famous report in 1967 under the very American title of “The Challenge of Crime in a Free Society”.

The committee also discussed the issue of benchmarks for measuring crime. It felt that the FBI index, the official police statistics of
Crime Surveys as Tools of Policy-making

In the United States, the accuracy of crime statistics was unsatisfactory and that surveys might provide a better measure of crime. After a couple of pilot surveys in several American cities, the Bureau of Justice Statistics developed, from 1973, a National Crime Survey that later was renamed the National Crime Victimisation Survey. At that time, subjective indicators were the big new fashion in social sciences. Rather than trying to measure difficult concepts, such as crime, inflation, housing or unemployment, by asking people what they thought about such things, monetary inflation, the value of real estate or state of the labour market were seen as better alternatives. Crime was an obvious candidate to be measured through subjective indicators. Rather than counting offences actually committed or reported, the public at large could be asked whether they had experienced crime or how safe they felt whilst walking in their neighbourhood.

Soon after the first pilot surveys conducted in American cities, Marshall B. Clinard (who died recently on 30 May, 2010, at the age of 98) developed in the seventies that surveys could become formidable tools for comparison of crime across international borders. With the crime victimisation surveys conducted in Zurich and in Stuttgart (Germany) in 1973, Clinard produced the first crime surveys outside the United States, and the first international comparison based on crime surveys to date. Indeed, international surveys require uniform offence definitions and standardised sampling and interview methods. Of course, supplementary questions are possible in all participating countries, but the key questions have to be identical. The importance of the standardisation of questionnaires and fieldwork has been shown by Richard Block, another American criminologist, during the 1980’s. He tried to compare crime rates across several countries by breaking down data from existing national surveys, which surveys had been conducted since the 1970’s and early 1980’s, in England, the Netherlands and in Switzerland. Indeed, this endeavour turned out to be relatively unsatisfactory and stimulated the search for truly international surveys. The realisation of such plans, however, needed the technical innovation of computer-assisted telephone interviews (CATI) which first became available in the 1980’s.

Innovative research techniques

This new technique was first used in Switzerland through a national survey. In 1984, Switzerland had, at this time, an unusually high telephone-per-household ratio of 98%, thus creating favourable conditions for survey companies to offer CATI surveys. This they began doing. Besides far lower costs (less than 15 percent of a traditional personal interview), CATI offers a series of advantages that are highly relevant for crime victimisation surveys. Probably the most important being the use of electronic questionnaires, which allow highly complex filters. If a respondent tells the interviewer that he or she was the victim of robbery, for example, he or she will go through perhaps 50 questions concerning stolen items, reports to insurance companies or to the police, the reaction of the police, whether there were injuries, the nature and treatment of any injuries etc. If the respondent did not experience robbery, all these questions will be skipped. In the case of a paper questionnaire, these questions would obviously
be printed, and the interviewer would have to turn dozens of pages to find the next appropriate question. The respondent might well be discouraged when he or she sees the interviewer opening a questionnaire as thick as a phone book. On the ‘phone, however, most interviews did not last more than 10 to 15 minutes, and only very few interviews (with seriously victimised respondents) lasted as much as one hour. The complexity of filters is therefore, a critical feature whenever the interviewer has to ask many questions to a few people (and few questions to many people), as typically in the case of crime surveys. CATI offers invaluable advantages in this respect.

Another advantage is anonymity. As studies on policies towards AIDS revealed (where very embarrassing questions had to be asked in surveys), CATI turned out to offer more anonymity to respondents compared with other interview methods. Meeting a subject personally means sacrificing some anonymity of the interaction, especially in smaller countries, since it is never certain that the interviewer will not sit on the same panel or meet in the same club or pub a few weeks later.

However, the critical advantage may still be costs. If less money is spent per interview, it is possible to provide a larger sample. Given the difference in price compared to face-to-face interviews, CATI allows the interviewing of six times more people. Assuming that serious crimes may hit perhaps one percent of the sample, in any given year, interviewing a sample of 1,000 people will mean locating an estimated 10 victims of serious assault. With such numbers, you cannot conduct sophisticated analyses. If a budget allows for interviewing 8,000 people, however, it will be easy to find 60 victims and be able to successfully refine the study. Therefore, lower costs per interview directly affect the scientific potential of a survey.

I am particularly pleased to be speaking about these issues here in Barcelona, as it was here that the international crime victimisation survey was designed, during a meeting of Jan van Dijk, Pat Mayhew and me with our Catalan friends in 1988. Given the past experience with the Swiss national crime surveys of 1984 and 1987, it was decided that we would adopt CATI as the interview method. With the low budgets that this method required, participating in the survey became possible for 11 countries in Europe, Australia, Canada and the United States.

**The method is less important than the way of asking questions**

After the publication of the key findings, the ICVS was heavily criticised. It was said that we investigated serious and eventually traumatic experiences of people through a market research instrument designed for the sale of shampoo or cosmetic products. Many people also felt that face-to-face interviews would have been much better suited for this theme. The point is, however, that if we had opted for face-to-face interviews rather than CATI, the international survey would never have got off the ground for budgetary reasons. Many countries could afford 30,000 €, but with a sum six times this budget, maybe 3 or 4 countries would have been able to participate. Alternative methods, such as mail questionnaires, have become outdated.
Crime Surveys as Tools of Policy-making

CATI was compared experimentally in the Netherlands, in Switzerland and in Germany with face-to-face surveys and mail questionnaires. In all instances, the common conclusion was that the interview method does not really matter. However, what is important is how questions are worded, although much less attention is usually given to the questionnaire's design. Particularly important is how events are located on a time-line. The usual way of asking “what did happen to you over the last 12 months” typically produces plenty of wrong temporal locations. It is far more precise to ask respondents to relate what has happened to them during the past five years (or what comes to their mind), and to ask in a second question more precisely when it actually happened (ie during the current year, the last year, the year before or longer ago). All these conclusions have recently been confirmed in connection with the International Self-reported Delinquency Project. Again, it is not the interview method (Internet vs. paper-pencil questionnaires) that matters, but the way questions (particularly those regarding temporal location) are phrased.

An interesting question is why Europeans (such as the designers of the Swedish and the French crime surveys) stick so much to the usual “last 12-months” formula. The reason may be the difficulty in departing from a standard pertaining in America, and the fact that the National Crime Victimisation Survey (NCVS) uses a similar format. The design of the American NCVS required, however, that the same household be interviewed every six months, i.e seven times in all over three and a half years. When visiting a household, the interviewer knows what the interviewee had reported during the preceding interviews and can, therefore, make sure that the newly reported incident is not a repeat of a previously reported incident. This design is extremely expensive and unaffordable in Europe. Being unable to follow the American model, we should obviously adapt our questionnaires to our surveys’ design. Outside the Netherlands and Switzerland, no-one has ever thought to query the efficacy of the flat “12-months formula” rigorously, despite the huge samples (and the considerable budgets) in England, Sweden and France. As the experiments conducted in the Netherlands by Annette Scherpenzeel suggest, the 12-months limit produces crime rates that are between 100 and 200 percent too high. There is good reason to presume that the French and Swedish survey rates are similarly inflated.

There are three more reasons for asking victims about incidents that are older than 12 months. Firstly, we do not find many victims of serious crimes in any given sample. If we ask about the last five years, we may have a much better base for statistical analyses. Secondly, it can be more than frustrating for the victim (and the interviewer) if a very serious crime cannot be taken into account because it happened shortly before the 12-months limit. Thirdly, policy-makers will certainly be interested in seeing how the reaction by the police and other bodies (such as victim support schemes) are evaluated by the victims. The successful prosecution of an offence usually takes more than twelve months. If we do not care about older events, policymakers simply will never get crucial information on how the victims feel about the way in which their case was dealt with.

The Barcelona initiative was a success story. In 1990, we started with 14 countries. In 1992, there were already twenty. In 1996, the questionnaire was revised and enlarged by including offences such
as corruption. The two studies of 2000 and 2005 followed the same line and covered in all roughly 60 nations. In 2010, given the boom of cell phones, CATI may no longer be the method of choice. With the increase in household Internet access, this medium will become a survey tool, and CATI will serve as an additional option for others. This new survey will be experimental in nature and limited to five EU countries (Netherlands, United Kingdom, Denmark, Germany and Poland) and Switzerland. Since the Swiss survey will officially be independent, we can add a number of additional experimental tests that may be useful for the future ICVS (or EU survey) model.

Measuring new offences through surveys

The ICVS was also innovative in measuring corruption. By asking respondents whether they had ever bribed an official (or were asked, by an official, for a bribe) during the last 5 years, and then more precisely when and under what circumstances, we were able to produce an interval scale of all participating countries. Transparency International, on the other hand, produces a rank order scale. Unfortunately, the two measures do not correlate as perfectly as one might expect because the Transparency International Index (TI) awards considerable weight to minimal differences (as we have them in Western Europe) that, on the ICVS scale, never would be considered as statistically significant, and reduces huge differences in the extent of corruption in other parts of the World to just a matter of a few ranks. A further problem is how TI measures corruption, namely through perceptions (or, if one wants to say it frankly, prejudice). I have tried many times but never succeeded in getting more information on what they actually measure. I believe the ICVS offers a far more reliable and better method of measuring in the future. It allows us to measure critical variables concerning possible causes of corruption, such as undue delays in the bureaucratic handling of legitimate citizen demands. For example, one might imagine an item in the questionnaire asking respondents how long it may take, in their country, to receive a new driver’s licence once one has lost it. Theft of handbags or purses with a driver’s licence can happen in any country around the world. Therefore, respondents all over the world would be in a situation to answer such a question, and the differences that are found might be revealing. Recently, I learned from an Italian student whose bag was stolen with the driver’s licence in it that it had taken nine months for him to get a duplicate – in a similar situation, it took me three days in my own country.

Statistics and surveys as trend indicators

Crime surveys are helpful also when it comes to assessing crime trends. Generally speaking, survey-based crime trends match trends in police-recorded crime rather well. Crime surveys allow an assessment of whether victims are increasingly inclined to report victimisations to the police. Indeed, this does not seem to be the case in most countries and in more substantial terms. Crime surveys as such also show, by the way, that – contrary to theft – violent crime has not
generally decreased in Europe, as many observers may believe. The American experience does not necessarily reach Europe with a delay of some years, as many may intuitively presume. There are, however, surprisingly varying trends between countries that, so far, are not really understood. As far as violence is concerned, the changes may be related to a revolution in leisure-time activities and other routine activities that have not so far been fully documented.

When we look at the match between police-recorded and survey-measured crime across countries, we find surprisingly good correlations for offences like burglary, theft of personal items and assault. I should add, however, that the police data are from the European Sourcebook of Crime and Criminal Justice Statistics (ESB) where considerable efforts were made to “streamline” definitions and counting conventions as far as possible. The exception is robbery, but this crime often concerns businesses that are not covered in crime surveys.

In sum, therefore, police-recorded data (as collected in the ESB) can compete with survey-data under certain conditions. Surveys may, despite these encouraging results, still be better at grasping international (or interregional) differences, whereas police-recorded crime may be more efficient at identifying crime trends. The reason is that counting conventions vary considerably from country to country, but much less so over time. The police and insurance companies dealing with injuries deal probably with more serious forms of violence, whereas surveys may be heavily biased towards more trivial forms of assault.

In any case, the question is not to know whether police or survey data should be preferred, but how survey-measures could be made more meaningful. My suggestion, therefore, is that we should collect more relevant independent variables to better understand why crime increases or decreases. It would be extremely helpful to have more data on lifestyle, routine activities, night life, drinking habits, drug abuse etc. Surveys can very well explain cross-national differences (or variations over time). To offer you an example, last night I was meeting a friend here in Barcelona for dinner. At around 11 PM, he told me that he would have to leave in about ten minutes because his last train (to reach his home in the suburbs) was departing in about 20 minutes. Tempi passati, in my country. Today, trains run in Zurich and other major Swiss urban areas throughout the night. That has dramatically changed life styles and, of course, produced an increase of violent crime. If one hundred thousand young people meet in the city centre between midnight and six in the morning, why should this leave crime unaffected? Unfortunately, we have so far, very little comparative data available that include these variables.

How to make crime surveys even more relevant

In future victimisation surveys, we should pay more attention to characteristics of offenders. It is easy to ask victims of contact crimes whether the offender was male or female, how old he or she might approximately have been, and whether the victim thinks he or she was of local or foreign origin. The International Crime Victimisation Survey never collected such data. If it had, as I suggested in 1990,
we would have a lot of crucial information available on the role of migration on crime beyond bald police statistics. Unfortunately, we missed this opportunity but we may still try in the future.

In sum, crime surveys are indispensable for policy makers to see where they stand, predict future trends and how things might be changed. In order to achieve these goals, it is important that crime surveys include many relevant independent variables, that they take into account past events (and not just victimisations experienced over the last 12 months), and that they can be repeated regularly (though not necessarily annually). In other words, it is crucial that they respect certain budgetary limits and that no single aspect is being favoured at the expense of others. For example, it would be most unfortunate to extend samples beyond 10,000 or 20,000 interviews, or to use expensive interview methods such as face-to-face interviews. Such choices, whatever statistical arguments may be advanced in their favour, will necessarily require crucial changes, such as the array of independent variables. That will leave policy-makers with highly sophisticated crime measures, but without explanatory variables that may help to understand differences, trends and changes. Policy-makers will hardly find much interest in such surveys – it is easy to guess that, at the next budgetary crisis and given their huge costs, they will be the easy targets for cuts.