FORECASTING CRIME RATES:
A REVIEW OF THE AVAILABLE METHODOLOGY

Dennis Weller

July 1979

Technical Report CERDCR-4-79

Prepared under Grant No. 77-NI-90-0071 from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice.

Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official positions or policies of the U.S. Department of Justice.

CENTER FOR ECONOMETRIC STUDIES OF THE JUSTICE SYSTEM
Hoover Institution, Stanford University

Prepared by the Center for the use of the United Nations Congress on the Prevention of Crime and Treatment of Offenders under the Center's Technical Assistance Program with the Department of Justice.

I. Introduction

In this paper we discuss a number of issues involved in forecasting the incidence of crime. We will review a variety of approaches which might be used to produce crime forecasts, and consider some of the literature dealing with those methods that have been used. We will evaluate the appropriateness of each of these approaches, considering the use to which the forecast is to be put and the availability of data. Finally, we will discuss some avenues of research that might be pursued to improve upon the existing models for crime prediction.

The literature of crime forecasting is not particularly rich. Researchers in the field of criminal justice have devoted more effort to the testing of hypotheses concerning the determinants of crime than they have to the prediction of future crime rates.

Yet forecasting models could be of service to criminal justice planners in a number of ways. These may be divided between the uses of short term and long term forecasts, and, further, between those models intended simply to predict the most likely values for crime rates in the future, and those designed to predict the effect of a change in some policy variable.

Long term forecasting is likely to be of greatest interest to planners at the state and national level, who seek to anticipate the demands of the...
II. Review of Forecasting Approaches

The incidence of crime is a difficult thing to measure, much less to predict. We will discuss some of the conceptual and practical difficulties in obtaining data on crime rates in Section III. In this section, to review existing forecasting methods, we will suppose that there is a relevant variable that we are able to define and measure in a satisfactory way. For purposes of discussion, we will call that variable the crime rate.

A. Extrapolation from Past Time Series of Crime Rates

The most straightforward approach to forecasting crime is to project the value of the crime rate into future periods by extrapolation from its values in prior periods. A number of techniques are available for this univariate approach; each assumes that there is some underlying pattern to the existing time series, together with a random element. These techniques include the following, in increasing order of complexity:

I. Moving Average. In this method, the predicted value is the simple average of an arbitrary number of previous observations. This is the simplest means of removing randomness from the data. The method of exponential smoothing is similar, except that more recent observations are weighted more heavily than older ones. Chamberlain (1971) employs exponential smoothing techniques to predict rates for specific crimes in Los Angeles precincts.

2. Regression. In its simplest form, this involves simply plotting a line through the available data points, with time as the explanatory variable and the crime rate as the dependent variable. A more sophis-
ticated method is the autocorrelation model, in which the explanatory variables are the lagged values of the crime rate in previous periods. Regression techniques may also be used to fit a moving average model. In this approach, the explanatory variables are the error terms, i.e., the difference between the predicted crime rate and the actual crime rate for each sample period.

3. Box-Jenkins and other more elaborate techniques. Various combinations of the methods mentioned above can be used to account for complex patterns in the data. There may be, for example, a general upward trend in the crime rate, around which seasonal swings occur.

Box and Jenkins (1970) have developed a general approach to these problems in which a well-defined procedure is used to select, fit, and verify the proper combination of autoregressive and moving average models. Deutsch (1978) uses a Box-Jenkins approach to model specific crime rates in ten major cities in the United States, using monthly data from 1966 to 1975.

These univariate projection models offer a number of advantages. They are conceptually simple, since they model the observed pattern in the data without attempting to represent the underlying mechanism which gives rise to that pattern. The only data required is a time series on the crime rate to be predicted. While the Box-Jenkins approach requires both greater skill and more computation than the simpler methods, it offers three compensating advantages. The first is accuracy. Given the current state of development of more elaborate behavioral models, and over a short time horizon (say, one or two years), the results of extrapolation from a Box-Jenkins model will probably be at least as accurate as those of any behavioral model. The second advantage is that of generality. The approach can be applied to virtually any kind of pattern in a time series. The third advantage is that the Box-Jenkins model offers not only a set of computations, but also a consistent set of criteria for deciding which computations to apply to the data.

The major disadvantage of the univariate approach is, in fact, the same as its major advantage. It is simple, which is another way of saying that it has no behavioral content. The model projects the observed patterns into the future without inquiring into the underlying factors which cause the patterns, or into the likelihood that these underlying factors will continue into the future as they have in the past.

Errors in forecasting can occur in two different ways. The first results from the random error occurring in any statistical model. The model does not fit the data perfectly; there will be error terms which represent variations not explained by the model. But it is possible to measure this error and to give estimates of the confidence with which the predictions are made. Depending on the type of model used and the nature of the data, the amount of random error may increase as the forecasting horizon is extended. Deutsch (1978), for example, computes confidence intervals for his projections of crime rates over a thirty-six month period; his first monthly projection predicts that the number of robberies in Cleveland in August 1975 will be 657. He specifies, further, that the actual observed value will fall, with a probability of .95, between a lower bound of 565 and an upper bound of 749. His estimate for the number of robberies in the thirty-sixth month is
The confidence interval for this estimate, however, is much wider: the lower bound is 481, and the upper bound is 1,244.

A potentially more serious source of error, however, is a shift in the underlying mechanism which generates the crime rate. This type of model makes no attempt to represent such a shift. While possible even with a short time horizon, changes in the underlying social or institutional structure should become more likely as we attempt to extend our predictions further into the future. By including some representation of the underlying structure in the model, though, it may be possible to predict at least some kinds of structural shifts.

A further limitation of extrapolation methods is that they are not capable of predicting the effects of alternative policies since there are no explanatory variables to be manipulated. These methods can be used, however, in interrupted time series experiments to judge the effectiveness of past changes in policy. This can be done by looking in the pattern of crime rates for a shift which can plausibly be supposed to result from the change in policy. This approach is used by Deutsch and Alt (1977) and by Hay and McCleary (1979) to evaluate the effectiveness of a gun control law in Boston. An alternative method is to estimate a prediction model using data from the period before the policy change, and then to use the model to predict what the crime rate would have been in subsequent periods if the new policy had not been in effect. Chamberlain (1971) employs this procedure to measure the effect on crime in Los Angeles of the introduction of helicopter patrols.

B. Indicator Tracking

A first step in modeling the underlying determinants of the crime rate is the identification of social or economic factors which appear to be related to the level of crime and which are subject to measurement. A simple search for "correlates" of crime, however, is likely to be even less edifying than past attempts to compile lists of "leading indicators" of economic activity. The mechanism by which the explanatory variables influence the crime rate must be specified before any attempt at prediction can be made.

1. Multiple Regression is an estimation technique which can be used to supply a quantitative structure to the relationship between the crime rate and its determinants. Brenner (1978), for example, has carried out a multiple regression analysis of the effect of unemployment, inflation, and per capita income on the rate of homicide in the United States. This involved fitting an equation in which the homicide rate was the dependent variable and the levels of unemployment, inflation, and income were the explanatory variables. Since the influence of changes in each of these three variables was supposed to be felt only gradually through time, lagged values of each variable were included. For each variable, a separate lag structure was estimated to specify the rate at which these delayed effects took place. Brenner's results suggest that a lower unemployment rate would lead to a lower homicide rate.

The advantage of the multiple regression approach is that it permits the inclusion of some behavioral content into the prediction model within a framework which is still conceptually (and computationally)
rather simple. It allows the researcher to measure the contribution of each explanatory variable while controlling for the other variables. To the extent that some of the explanatory variables are thought to be determined by criminal justice policy, the effect of a change in policy on future crime rates can be predicted.

The disadvantages of multiple regression methods arise in part because a single linear equation cannot adequately represent the underlying relationships, which may be quite complex. The explanatory variables may, for example, influence one another (as in the case of income, unemployment, and inflation, for example). These interactions cannot be measured in a single regression equation.

Since series are needed for the explanatory variables, multiple regression models require more data than do univariate methods. Further, if predictions are to be made, then forecasts of the explanatory variables are required. If the explanatory variables are themselves more difficult to predict than the crime rate, the usefulness of the model will be limited. It will still be possible to perform policy analysis with the model by computing the effect of a shift in one of the explanatory variables on future values of the crime rate. This is what Brenner has done in his measurement of the effect of unemployment on crime. However, unless the variable under consideration is directly under the control of the policy maker, and unless it is independent of the other explanatory variables, the validity of this exercise is open to question. The policies which might be adopted to reduce unemployment by a given percentage will surely have an effect upon inflation and income as well.

These interactions are not accounted for by the model, nor are any other mechanisms by which the policy measures might possibly influence the crime rate.

The selection of appropriate explanatory variables for inclusion in a regression model is a matter of some difficulty. Simply searching a large list of socioeconomic indicators for those with the highest correlations may produce some undesirable results. If, for example, an important variable is left out of the model, the results will be biased. Bold (1978), for example, has shown that the inclusion of some plausible explanatory variables—such as the age distribution—into Brenner's analysis changes the results considerably, eliminating the previously observed effect of unemployment on crime. Conversely, if a variable is included without sufficient theoretical justification, its correlation with the crime rate during the sample period may prove to be a spurious one.

2. The use of Delphi Techniques constitutes an admission that no adequate quantitative model can be constructed to address the problem at hand. This approach, developed by Olaf Helmer (1966) at the Rand Corporation, seeks to substitute, instead, the informed but subjective judgment of a panel of experts on the likelihood of various events in the future. Delphi, then, is a set of procedures for soliciting the opinions of the experts, tabulating the results, and feeding information back to each participant so as to encourage the convergence of opinion toward a consensus.

While it is true that Delphi methods can be used to predict events when it is not possible to construct an analytical model, it is dif-
ficult to know how much confidence to place in a prediction of this kind. There is no objective means for evaluating its accuracy, or for replicating its results. Difficulties arise in deciding who is an "expert" on a particular topic, and in formulating the questions to be put to the panel. At best, it would seem that Delphi might prove useful in identifying possible variables for inclusion in analytical models.

C. System Flow Models and Microsimulation

One approach to modeling the interactions of different elements in the criminal justice system has drawn upon the techniques of operations research. Most applications of this approach have involved the construction of flow models to trace the movement of offenders through a branching process from arrest through trial to imprisonment. Some model builders have estimated the probabilities involved at each branch in the system; others have allowed the user to specify hypothetical values for the purposes of evaluating different policies. These models have not, in general, attempted to model the generation of crimes or arrests; rather, they have sought to predict the stocks and flows in the justice system, given the number of offenders entering the system.

There have been a few efforts, however, to predict crime or arrest rates within the context of operations research methodology. For example, Deutsch, Jarvis and Parker (1979) have extended the work of Deutsch and Alt (1977), previously mentioned, through the use of a network flow model to predict the effect of crime control policy on the displacement of crimes among locations in a metropolitan area.

Blumstein, Cohen and Miller (1978) have developed a model to predict prison populations. Unlike most such models, theirs attempts to predict the flow into the justice system, in the form of arrests, as well as the path followed by offenders through the system. Arrests are predicted by disaggregating the population into cells defined by demographic characteristics such as age, race, and sex. A linear trend projection is made of the probability of arrest for the members of each cell. The authors feel more confident in using projections of demographic characteristics of the population as explanatory variables than they would in using predictions of such variables as unemployment. The projected demographic-specific arrest rates can be translated into crime rates if we are willing to assume that the ratio of crimes to arrests will remain stable. The difficulty of using such a disaggregated model to predict crime rates directly, of course, is the lack of information on the demographic characteristics of criminals in general, since we are only able to observe those criminals who are arrested.

Demographic projections tend to be quite accurate in the short run, since changes in demographic patterns take place slowly. In the very long run, however, they are subject to changes in behavior, such as migration and family formation, which can upset the predictions. As in other types of forecasting, some behavioral content must be incorporated into the model if such changes are to be anticipated. One development in this direction is the use of microsimulation models, which attempt to predict the life paths of a large sample of individuals drawn at random from the population. Some of these models are extremely elaborate; the DYNASIM
Fox (1978) has constructed a simultaneous equation model for the purpose of predicting the national crime rate (FBI index crimes per 100,000 population) in the United States. The model was estimated using annual data for the years 1950 to 1974. Table 1 presents the forecasts generated by Fox's model for the years 1975 to 2000.

Another model has been formulated under the sponsorship of the National Institute of Law Enforcement and Criminal Justice (1978) for use in conjunction with the National Manpower Survey. The main purpose of this model is to predict the demands of the criminal justice system for different kinds of labor. For this reason, the section of the model which describes the production functions of the various criminal justice subsystems (police, courts, corrections) is relatively elaborate. The model also includes, however, an equation which estimates the crime rate. The data used to estimate the model was a pooled cross section and time series including information by state for the years 1970 to 1974.

The simultaneous equation approach offers a number of advantages. It permits the consideration of interactions among the various explanatory variables—something which, as we have pointed out, a single equation model cannot do. In this way, more complex theoretical hypotheses can be modeled. Further, it may be possible to enhance the usefulness of the model for policy analysis by specifying, within the model, the relationship between policy variables and other variables which more directly affect the crime rate. For example, rather than considering the effect of an arbitrary increase in arrests, we can consider the effect of an increase in expenditures for police on the crime rates through its effect on arrests (and perhaps on

model created at the Urban Institute is a notable example. A large number of economic and social characteristics are considered in this model, and the complexity of its structure allows the effects of these variables to interact in a number of ways. This not only improves the reliability (and the specificity) of the demographic predictions; it also permits the model to address matters relating to policy, such as future rates of saving and labor force participation.

There are two obstacles to the utilization of large scale microsimulation models in the prediction of crime rates. One is the very high cost of constructing and operating a model of this size. The other is the aforementioned difficulty of obtaining disaggregated information on criminal behavior. It may be possible to combine the microsimulation of some aspects of the model with a simultaneous equation approach to other sectors. Something of this sort is, in fact, done in DYNASIM, since a macroeconomic model is appended to it.

D. Simultaneous Equation Systems

Perhaps the most promising approach to overcoming the limitations of the multiple regression models discussed above is the combination of several regression equations into a simultaneous system. Each equation may be used to specify the hypothesized behavior of one part of the system—the effect of police expenditures and other factors on arrests rates, for example. This arrangement affords a much greater opportunity to incorporate some theory of behavior into the model. The most widely known application of simultaneous equation systems has been the development of econometric models for use in evaluating monetary and fiscal policy.
Table 1

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Crime Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>6343.2</td>
</tr>
<tr>
<td>1976</td>
<td>6449.7</td>
</tr>
<tr>
<td>1977</td>
<td>6512.3</td>
</tr>
<tr>
<td>1978</td>
<td>7053.0</td>
</tr>
<tr>
<td>1979</td>
<td>7255.3</td>
</tr>
<tr>
<td>1980</td>
<td>7302.0</td>
</tr>
<tr>
<td>1981</td>
<td>7287.5</td>
</tr>
<tr>
<td>1982</td>
<td>7282.8</td>
</tr>
<tr>
<td>1983</td>
<td>7501.8</td>
</tr>
<tr>
<td>1984</td>
<td>7507.8</td>
</tr>
<tr>
<td>1985</td>
<td>7516.9</td>
</tr>
<tr>
<td>1986</td>
<td>7650.2</td>
</tr>
<tr>
<td>1987</td>
<td>7732.1</td>
</tr>
<tr>
<td>1988</td>
<td>7772.8</td>
</tr>
<tr>
<td>1989</td>
<td>7805.9</td>
</tr>
<tr>
<td>1990</td>
<td>7984.8</td>
</tr>
<tr>
<td>1991</td>
<td>8115.7</td>
</tr>
<tr>
<td>1992</td>
<td>8468.2</td>
</tr>
<tr>
<td>1993</td>
<td>8717.0</td>
</tr>
<tr>
<td>1994</td>
<td>8911.2</td>
</tr>
<tr>
<td>1995</td>
<td>9107.3</td>
</tr>
<tr>
<td>1996</td>
<td>9290.1</td>
</tr>
<tr>
<td>1997</td>
<td>9421.2</td>
</tr>
<tr>
<td>1998</td>
<td>9536.6</td>
</tr>
<tr>
<td>1999</td>
<td>9629.0</td>
</tr>
<tr>
<td>2000</td>
<td>9715.4</td>
</tr>
</tbody>
</table>


The relationship between police expenditures and crime is a frequently cited example of another econometric problem, that of separating the effects of two variables upon one another. The expenditure on police may plausibly be thought to be a determinant of the crime rate. It is equally reasonable, however, to suppose that the level of crime is a determinant of the public's willingness to pay for police protection. The problem may be dealt with more readily in a multiple equation system. Fox, for example, uses a set of lagged values of crime rates to estimate their effect on police expenditures over time as an aid to establishing the direction of causality.

The disadvantages of the simultaneous equation method include increased complexity, computational difficulty, and in most cases, an increased demand for data. At this level of complexity, considerable care is required to specify a model that is theoretically sound and that also satisfies all of the assumptions necessary to implement and estimate a system of equations. The previously noted difficulty of obtaining forecasts of the exogenous, or predetermined, explanatory variables, is also present here, although an effort can be made to limit the number of these variables that are not subject to policy. Fox, for example, uses a prediction of the Consumer Price Index (CPI) supplied by a prominent econometric forecast. The rate of inflation implied by this series for the year 1979 is 5.1%. (This demonstrates, by the way, if demonstration is needed, that macro econometric models are themselves less than perfect forecasting tools).

While the simultaneous approach offers substantial scope for expanding the behavioral content of the model, there is no guarantee that the theories
incorporated into any given model will be particularly powerful. Although the Fox model is carefully estimated, its theoretical underpinnings are rather weak. The only "economic" variable included, for example, is the absolute level of the CPI, which would not appear, on the face of it, to have anything to do with the level of crime.

An examination of the forecast in Table 1 reveals that the predicted crime rate has an upward trend, surrounded by a cyclical swing. As Fox points out, the upward trend is provided in the model by the CPI, while the cyclical movement comes from changes in a demographic variable—the proportion of nonwhite youths in the population. The growth in the crime rate moderates in the 1980s as the "baby boom" cohort ages, then increases in the 1990s as their children become young adults.

III. Problems of Definition and Measurement

In previous sections we have assumed that there is a well-defined quantity—the crime rate—for which accurate measurements have been collected in the past. The only difficulty was in the method by which future values could be forecast.

In practice, however, the most widely used measures of crime are subject to considerable error, even in relatively developed countries. In the United States the standard crime statistics are the Uniform Crime Reports (UCR), but many law enforcement agencies do not report to the FBI all of the data requested by the UCR.

Official crime records, no matter how complete, include only those offenses which are reported to the police. Results from victimization surveys suggest that a great many crimes are never reported. These surveys provide an alternative method of measuring crimes, but they unfortunately involve conceptual and methodological difficulties of their own. They have also been quite expensive. No other survey program in the United States, except the decennial census, has a larger sample size. A discussion of the limitations of the UCR and victimization data is given in Skogan (1976).

Crime data is not unique, of course, in being difficult to collect. Most aggregate measures of economic and social variables are based on imperfect data. Unemployment, for example, is a very difficult quantity to define and measure. If these data collection difficulties arise in more developed countries, it is only to be expected that accurate information should be even harder to collect in developing countries. This must necessarily limit the nature of the models that can be employed in those countries. This is unfor-
tunate, since the need to account for changes in underlying social and economic patterns would appear to be greatest in these rapidly developing countries.

If the difficulties of collecting accurate and consistent data on crime within each country are formidable, those involved in building a set of data comparable across countries appear insurmountable. It is difficult to see how the results of the cross-sectional studies which would utilize such a data set could justify the expense of assembling it, in view of the innumerable factors for which such studies would have to control.

Summary and Assessment of Future Development

Based on our review of the available techniques for the forecasting of crime, it is possible to make some general recommendations as to the most suitable method, depending on the use of the forecast and availability of data and other resources.

First, it would appear that extrapolation by means of one of the univariate methods would be most appropriate for generating short-term forecasts at a reasonable cost. It is doubtful that a more complex behavioral model will be any more accurate in predicting one or two years ahead. A more complex model will be needed, however, if projections of the effects of alternative values of some policy variable are to be produced.

As the time horizon over which the forecast is to be made increases, so do the dangers of pure extrapolation. The only hope for generating useful long-term (say, twenty-year) predictions lies in the development of a model which incorporates a reasonable set of theoretical hypotheses about the underlying social and economic mechanisms which influence the crime rate. Every forecasting model, of course, involves extrapolation, and no model can hope to anticipate structural changes perfectly. The use of a coherent theoretical model, however, offers at least the possibility of long-term forecasts in approximate terms and in the absence of major, unanticipated shocks.

Unfortunately, no such theoretical model is now available. Several elements that could be incorporated into such a model have appeared in the literature in recent years. A simultaneous equation system similar in form to econometric macro models would seem to be the most likely vehicle for
long term forecasting, although a large scale microsimulation model might also be used.

Areas in which the theory might be extended and developed for inclusion in a forecasting model include:

1) Definition and measurement of the risks and returns of crime for individuals, and the way in which they vary according to the characteristics of the individual, such as age, race, and education. The risks, of course, include the probability of arrest and conviction. The returns include the level of sanctions and, for property crimes, the expected gain for a crime of a given type.

2) Investigation of the risks and returns of alternative activities, such as legal employment. This would include the wage rate and the probability of being unemployed, according again, to individual characteristics.

3) Specification of the "production" process for various elements of the criminal justice system, such as police, courts, and correctional agencies.

4) Investigation of social and environmental variables which might help to explain different preferences, with respect to legal and illegal activities, across individuals, given a set of opportunities.

The combination of some or all of these elements into a workable model is something which is not likely to come about easily or quickly. Many of the topics listed above would occasion considerable debate among researchers in the field. The effect of sanctions on the crime rate, for example, is something which has yet to be precisely estimated. In addition, the problem of obtaining forecasts of the explanatory variables will continue to limit the scope of these models. Nonetheless, if accurate long term forecasts are felt to be useful, then the development of an accurate underlying theory of criminal behavior is the avenue of research which is most likely to contribute to achieving that goal.
NOTES


2 This does not guarantee, however, that two independent investigators will necessarily agree on the exact specification of a Box-Jenkins model. Compare, for example, the work of Deutsch and Alt with that of Hay and McCleary on the subject of gun-related offenses in Boston.

3 Probably the best known model of this type is JUSSIM, created by Belkin, Blumstein, and Glass (1972).


References


