Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Discussion Papers   Economic Growth Center

3-1-1990

Social Osmosis and Patterns of Crime: A Dynamic Economic Analysis

Raaj K. Sah

Follow this and additional works at: https://elischolar.library.yale.edu/egcenter-discussion-paper-series

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut   06520

CENTER DISCUSSION PAPER NO. 609

SOCIAL OSMOSIS AND PATTERNS OF CRIME:

A DYNAMIC ECONOMIC ANALYSIS

Raaj K. Sah

Yale University

March 1990

SOCIAL OSMOSIS AND PATTERNS OF CRIME:

A DYNAMIC ECONOMIC ANALYSIS

Raaj K. Sah, Yale University

Revised: March 1990

Abstract: Though crime and the fear of crime have a deep negative impact on the well-being of individuals and societies, there are several aspects of this phenomenon that are not adequately understood at present within economic models. For instance, individuals' perceptions of the probability of punishment (for a given crime) differ widely within and across societal groups, and these perceptions in turn differ from the actual probability of punishment. Also, even after controlling for various economic and deterrence variables, crime rates in different societal groups have often been observed to be strongly correlated with background variables like location, age structure and ethnicity.

One of the points of departure of this paper concerns the treatment of individuals' perceptions. In the literature based on Becker's seminal study, an individual's perceived probability of punishment is typically treated as an exogenous parameter that is the same for all individuals and equal to the actual probability of punishment. The present paper begins by developing a model in which an individual's perceptions and the resulting choice are endogenously determined. This model deals with some basic economic features of the information available to individuals, with how this information is generated within the economy, and with some of the consequences of the endogeneity of their perceptions.

This model of individual perception and choice, when aggregated to the economy-wide level, yields dynamic relationships linking certain features of the economy (including past crime rates), to individuals' current choices, and then to the current crime rates in different societal groups. These relationships suggest a possible way of understanding some observed patterns. They also suggest insights concerning, for instance, how criminality might evolve over time in a society, why two societal groups with roughly similar economic parameters might exhibit different crime rates, and why crime might spill over across different societal groups.

I. INTRODUCTION AND OVERVIEW

Crime rates in different societal groups are often strongly correlated with such background variables as location (city-centers versus suburbs versus rural communities), age structure, and ethnicity (blacks versus whites in the U.S., for instance).[1] As an example, a young black man in New York City is believed to be over ten times more likely than his non-black counterpart to commit a robbery. Empirical work has established such correlations *after* controlling for a large number of economic variables (among them, individuals' income, education, work experience, and intra-group income distribution) and deterrence variables (such as the penalty schedule and the resources spent on the criminal apprehension system[2]).[3]

Crime and the fear of crime have a deep negative impact on the well-being of individuals and societies. For instance, in the U.S., opinion surveys conducted during the last decade show that most people view crime to be the single most important problem they face, more important than the possibility of various economic crises or of military conflict. It is important, therefore, to attempt to improve the positive understanding of observed patterns including those noted above. In turn, such an understanding is essential for policy debates, regardless of the normative views one might adopt.

To see one of the points of departure of this paper, consider Becker's seminal study, which has inspired a valuable literature.[4] In this study, a key ingredient in an individual's choice of whether or not to be a criminal is his or her perceived probability of punishment. For brevity of discussion, we use the symbol p to refer to this probability (which represents an individual's *perception*), and use r to refer to the actual probability of punishment (which represents the corresponding *reality*). The model of the individual on which this literature is based assumes that p is an exogenous parameter. It abstracts from such questions as, what the determinants of individuals' p's are, which economic variables influence them and why, and what the consequences of these influences are. In fact, this literature is typically based on the much stronger assumptions that the p's are the same for all individuals, and that the common value of p is the same as r. That is, all individuals have identical and exogenously given perceptions, and these are the same as the

---

[1] See Bureau of Justice Statistics (1983, 1988) for exhibits of U.S. data.

[2] For brevity, we use the phrase "criminal apprehension system" for various public activities to apprehend and punish criminals.

[3] See Pyle (1983, Ch. 3) for a review of more than twenty empirical studies and for references to other studies and reviews.

[4] Becker (1968). See Andreano and Siegfried (1980), Becker and Landes (1974), Heineke (1978), Klevorick (1985), Pyle (1983, Chs. 2 and 5), Schmidt and Witte (1984), and references therein, for reviews, applications, and many extensions. While these and other extensions are important, they are not central to the focus of the present paper.

reality.

It is useful to modify these simplifying assumptions in the light of the evidence concerning individuals' self-reported perceptions that has accumulated, based on surveys conducted during the last two decades. Among the findings that are important for the present paper are the following:

(a) A large variance in p's across individuals within a societal group is observed, regardless of the background characteristics of the group (for example, income, education, ethnicity, etc.), and regardless of how narrowly defined the category of crime is under consideration (for example, theft of $50). This finding has been reported in many studies, including those cited below.

(b) There are significant differences in the distribution of p's across different societal groups. See Carter and Hill (1978), Piliavin *et al.* (1986), Richards and Tittle (1982) and references therein.

(c) An individual's p changes over time. See Piliavin *et al.* (1986) and other longitudinal studies cited therein. The mass media does not significantly influence individuals' p's, whereas past experiences of one's own and of one's acquaintances (for example, friends, peers, relatives) have a strong influence. The preceding two types of past experiences (that is, one's own and of one's acquaintances) do not appear to have significantly different influences on the p's. See Parker and Grasmick (1979).

The research currently available on individuals' perceptions is limited, perhaps in part because of a lack of adequate theoretical perspectives, especially economic ones, on these matters. In any event, evidence of the kind described above clearly suggests a need for thinking about the determinants of individuals' perceptions, no matter which model we use to understand how these perceptions get translated into individuals' choices. This is especially important for analyses based on economic models because they typically predict that an individual with a higher p will have a higher propensity for crime. This link, apart from being well-grounded in theory, has been empirically established in many studies based on individual-level data, including an exhaustive study by Montmarquette and Nerlove (1985).

As a step towards understanding ordinary individuals' perceptions, it is useful to note that even the experts who conduct serious empirical research on crime disagree on their estimates of the values of r in the past periods, let alone the current value of r. This well-known problem arises because it is difficult to ascertain the number of crimes committed. Crime statistics are inherently unreliable, since individuals and agencies have incentives, or the lack thereof, that cause incomplete or improper crime reporting. Broad indications of how flawed crime reporting might be are provided by studies of self-reported criminality. For example, in a sample of 2510 adult American males, 44% admitted to shoplifting, whereas the fraction

of the population punished for this crime is minuscule in comparison (O'Donnell *et al.* (1977)). Another study found that only 5% of the self-reported larcenies in Finland were ever detected by police (Anttila and Jaakkola (1966)). The informational issue we wish to emphasize here is that even though a researcher can, in principle, obtain more reliable data on the number of criminals who have been punished, the assessment of r remains a problem of costly statistical inference. It is also reasonable to posit that, compared to an expert researcher, an ordinary individual faces a much more difficult and costly inference problem.

Adopting the perspective that an individual's perception is an endogenous outcome of the nature of the information available to him, one may attempt to understand it in the following manner. Though an individual may receive raw data from several sources (such as his own past experience, past experiences of his acquaintances, and from crime reporting by the mass media), each source has its own costs, inaccuracies, randomness, and limited information content. For example, casual or indirect contact, or hearsay, yields unreliable data. An individual's observations from his own past experiences and from those of his acquaintances are more reliable, but they provide limited information because only a limited number of such observations are available (put differently, an individual faces increasingly higher costs of gathering, beyond some point, each additional observation of this type). Also, note that observations from the past experiences of one's acquaintances are not error-free because one needs to adjust them for the differences in abilities (for instance, the ability to escape punishment) between oneself and others. On the other hand, crime reporting by the mass media is an inexpensive source of data which is economy-wide (that is, it is common to many individuals), but it is largely irrelevant for the issue at hand. This is because intensive but partial coverage of a small subset of reported crimes, which might be chosen for its sensational impact, contains virtually no useful information as far as r is concerned. Moreover, there are no other institutions through which an individual can ascertain r.

Thus, from an economic point of view, some of the key features of the nature of the information available to an individual are as follows. (i) The information is limited (this reflects the cost of information as well as its unreliability). If this were not the case, then, contrary to the evidence, all individuals would know r. (ii) A substantial part of the relevant information comes to an individual from his "vicinity" (that is, oneself and one's acquaintances). (iii) To an extent, the individual's current information reflects, in a stochastic manner, the values of r in some of the past periods. This is because the past punishment experiences of those in one's vicinity depend stochastically on the values of r in the corresponding past periods. The past values of r may similarly affect other sources of data available to an individual, to the

extent that these sources contain any relevant information. (iv) The preceding information concerning past values of r is an imperfect predictor of the current value of r not only because it is local, limited and stochastic, but also because the value of r may be changing for a variety of economy-wide reasons. On the other hand, as the arguments presented above suggest, the individual has virtually no relevant economy-wide information.

We have attempted to capture these features in some simple models. These models can be made more realistic and elaborate in many ways discussed in the next section. However, it is important to note that our analysis draws upon these key features; the models we have used are only convenient ways to structure them.

A natural consequence of the features described above is that an individual's p may differ from r, and that p's may differ across individuals. This is consistent with the evidence noted earlier. Also, an individual's p may be influenced, in a stochastic manner, by such variables as past crime rates (because these may influence how many criminals one might observe in one's vicinity) and the values of r in the past (because these may influence how many criminals one might observe in one's vicinity who have been punished). The resulting perceptions of an individual, in combination with his opportunities from criminal versus non-criminal activities, determine his current choice of whether or not to be a criminal.

We move from an individual's choice to the economy-wide crime rate (which is simply the fraction of the individuals in the economy who have chosen to be criminals in a particular period) as follows. An individual's choice described above has a dynamic element because his current p is influenced by such variables as past crime rates and the past efficacy of the criminal apprehension system. These past variables, combined with the parameters currently faced by the individual, determine his current choice. These current choices, aggregated across individuals, determine the current crime rate. In turn, this crime rate affects the actual probability of punishment, r, in the current period. This is because for any given public expenditure on the criminal apprehension system, a larger number of criminals (equivalently, a higher crime rate) leads to fewer resources being spent on apprehending each criminal, thereby lowering r. The current value of r and the current crime rate, in turn, may influence some future crime rates.

To describe these dynamics, we use an overlapping-generations framework, in which a new cohort of individuals enters different societal groups in each period, while an old cohort leaves the economy. For individuals within a societal group and a cohort, the available information (which, it should be recalled, is an imperfect predictor of the current value of r) may become less divergent if the cohort is older. On

the other hand, this framework also allows for the possibility that perceptions can differ across individuals (especially in the younger cohorts) and that the distribution of perceptions can differ across cohorts and societal groups.

To recapitulate, then, we begin with a model of an individual's perceptions and choice which is based on reasonable assumptions, and the implications of which are consistent with the available evidence concerning individuals' perceptions. Using this as the starting point, we derive dynamic relationships, linking the features of the economy (including past crime rates and r's), to individuals' perceptions (which are intermediated by the nature of the information available to individuals) and choices, and then to the crime rates in different societal groups. Among the qualitative insights suggested by these relationships are the following:

(i) A long tradition of social thought has maintained that an individual's environment influences his propensity for crime.[5] Such views can be assessed by using models such as those in this paper, in which the environment plays an explicit role. Such an assessment has potential implications for the public debate on crime. For instance, some social theorists have contended that punishment does not deter crime because its "root cause" is the environment (see Sowell (1980, Ch. 9) for a review and criticism of such views). Though this contention is influential, it is not supported by the present analysis, which shows that, although the environment matters, the current crime rate will be lower if apprehension and punishment have been more efficacious in the past.

(ii) A society's past exerts economic influences. For instance, the current crime rate is higher if past crime rates were higher. Consequently, two societies can have different current crime rates, even if they currently have similar parameters, provided their past crime rates were different. Although past crime breeds future crime, it does not mean that a society is eventually doomed to complete criminality. The fact that new cohorts routinely enter the society and old ones leave plays a potential balancing role.

(iii) Our analysis has possible implications for the public debate on crime which differ from those that follow from analyses of crime based on a static model of individual choice. For example, since the latter do not deal with the consequences of past variables, they may overstate the effects of current policy variables (regarding deterrence and public expenditure) on current crime rates. Given past variables, current crime rates may be rather insensitive to current policy variables. By the same token, these analyses may

---

[5]An early public recognition of this view can be found in French crime statistics (Government of France, various years), which, since 1825, have gathered details on criminals' "milieu social."

understate the effects of current policy variables on future crime rates.

(iv) In order to understand some possible effects of the economy's parameters on its crime rates, we assess the impact of a change in such parameters as income distribution, individuals' initial beliefs, and parameters affecting the nature of the information available to individuals. As an illustration, if more individuals initially believe that they can get away with crime, then future crime rates will be higher, the actual probability of punishment will be lower, and more people will actually get away with crime.

(v) Intergroup differences in crime rates are a subject of ongoing debate. Our analysis suggests that significant differences among different groups' crime rates can coexist with relatively modest differences in their economic parameters, even if all groups face the same r. To see one of the reasons for this, recall the earlier discussion of the key features of the information available to individuals. The economy-wide sources of common data, such as crime reporting by the mass media, may be irrelevant in determining an individual's p, while the observations that he gathers from his vicinity may be far more important. The latter observations are likely to be more affected by the variables pertaining to one's own group than by those pertaining to other groups. The resulting differences in the distribution of p's across groups may contribute to differences in their crime rates.

(vi) Crime can spill over across groups. For instance, if a change in the parameters faced by a particular group (say, an increase in the pay-off from crime) raises its crime rate, another group's crime rate may also rise even though the latter group does not face any parameter changes. Such inter-group influences arise, in part, because of the impact of one group's variables on the perceptions of individuals in other groups, and on such variables as r, which are determined at the economy-wide level. Further, some societal changes, such as desegregation, may alter the exposure that the members of different social groups have to one another. In turn, this may affect the nature of the information available to individuals in different groups. Using this link, we analyze some possible effects of desegregation on different groups' crime rates.

To keep the length of the paper manageable, we provide only the outline of most mathematical derivations. For the same reason, as well as to focus on the newer issues analyzed in this paper, we adopt several boundaries, including the following:

First, the specification of individuals' choices is kept as sparse as possible. An individual's choice in each period is whether or not to be a criminal. However, our analysis applies to other specifications as well, such as those in which an individual chooses from a range of criminal and non-criminal activities that

can be undertaken with varying levels of effort within the same duration.

Second, the paper is restricted to a positive analysis in which the penalty schedule and the public expenditure on the criminal apprehension system are treated as exogenous variables. One of the objectives of the analysis, then, is to ascertain how crime patterns might change if these variables were altered. This analysis can be extended to incorporate an endogenous government response. For this, one would need to specify a theory of state (for example, the median voter model, a model of political pressure groups (see Becker (1983)), or a social welfare function) and an endogenous government budget constraint. The aggregate crime rates that will be observed in an economy will obviously depend, in part, on the nature of the government response. For example, if the government response is weak, then the outcomes will be similar to those in which this response is treated as fixed. On the other hand, the outcomes could be quite different if this is not the case. However, an analysis of endogenous government responses arising from alternative theories of state would make the present paper too long and unwieldy. Thus, it is best to treat such an analysis as a future topic, and to follow the standard practice of first dealing with the analysis in which the government is not endogenized. The analysis presented here is, in any case, an essential step in studying this topic.

Third, we have endogenized individuals' perceptions concerning r, but not their other perceptions. There is some evidence that individuals face an uncertainty concerning the penalty schedule. We have included this uncertainty in our description of individuals' pay-offs from crime versus non-crime activities, but have not endogenized their knowledge concerning pay-offs. This is partly for brevity, but also because our emphasis on the perceptions concerning r reflects an attempt to deal with a set of issues which have been viewed as important in the empirical literature on criminality. Moreover, there are significant differences between different categories of perceptions. Individuals' perceptions concerning penalties do not, by themselves, alter the actual penalties; they are altered by changes in the law. On the other hand, given a set of laws, the actual probability of punishment, r, is determined endogenously, partly as a consequence of individuals' perceptions concerning r. The endogeneity of r plays an important role in our analysis.

Fourth, environmental influences of the kind emphasized here may play a role in some other economic contexts. This paper focuses solely on crime because it is, in itself, an important social issue, and also because the available evidence concerning individuals' perceptions suggests a need to think about their determinants and consequences. Moreover, the analysis in other contexts could be quite different from

the one presented here. This is because the analysis would have to be based on context-specific issues and questions such as, what the environmental influence is, why it might be important, who the key economic actors are, what the relevant institutions are, and what specific questions are to be examined. (An example of an institutional difference is that, in the context of employment decisions, though a worker may not fully know the probability of his lay-off, its consequences are altered to some degree by explicit or implicit contracts or insurance; the corresponding issues are likely to be less important in the context of crime.) A discussion of such differences and similarities across different economic contexts is too broad a topic to be treated within the scope of the present paper.

Another aspect of our analysis that is worth highlighting is as follows.[6] We derive the economy-wide crime rate in different groups as a function of past variables and current parameters. This derivation is based on an explicit model of perception and choice of heterogeneous individuals. This dynamic relationship then forms a basis for studying the properties of the crime rates in different groups. To simplify some of the analysis (particularly, the impact of a change in the parameters), we focus on the steady-state crime rates (that is, the hypothetical situation in which the period-to-period changes in the crime rates are negligible). As in many other contexts of economic analysis, the steady-states merely provide a convenient apparatus with which to study the underlying economic forces. It is not being assumed or suggested that a real economy actually arrives in a situation in which period-to-period changes in the crime rate are zero; this is because numerous shocks and parameter changes impinge upon a real economy in a routine manner. Also, a steady-state does not depict any kind of equilibrium, because there are no agents in our economy who have the incentive or the capability to eliminate period-to-period changes.

A related issue is that multiple steady-state crime rates arise in our analysis as a consequence of the underlying dynamic processes (for example, the effect of past crime rates). This multiplicity of steady-states is different from the multiplicity of equilibria in some other economic models. As an illustration, consider the following adaptation of the standard Nash model. An individual first computes r in each of the numerous hypothetical scenarios in which the rest of the population makes different possible choices, and then determines his response in each case. (For example, he may choose to be a criminal in the hypothetical case in which most other individuals are assumed to choose to be criminals, because r will then be small.) This model will typically have multiple Nash equilibria. However, there are many reasons

---

[6] I have benefited from a discussion with Gary Becker on the material presented below.

why such a model may not be an appropriate vehicle for studying the concerns that underlie the present paper. First, the strategic assumptions behind this model are perhaps not best suited for the economies under consideration that have a large number of individuals (see footnote 7 below). Second, the assumption that each individual has the information to calculate r's for numerous hypothetical scenarios is inconsistent with the evidence and arguments presented earlier. Third, several questions and outcomes that our analysis can deal with are not satisfactorily analyzed within the Nash model. An example is the outcome that different groups with similar economic parameters can have different distributions of perceptions and different crime rates, even though they face the same r. Fourth, a multiplicity of Nash equilibria does not typically provide an explicit reason, within the model itself, for why an economy might be in one equilibrium versus another. In contrast, our analysis provides a reason for why two economies with similar current parameters may have different current crime rates. The presence of explicit historicity in our analysis also makes it possible to understand the process through which the effects of different parameter changes are felt.

The paper is organized as follows. Section II presents a model of individual perception and choice, and then derives some of the properties of an individual's propensity for crime. Section III describes the economy-wide crime rate, identifies some of its properties, and examines how it might change if some of the parameters of the economy change. Section IV presents a useful simplification in which individuals' perceptions are described by Bayesian inference. A relationship between the nature of the information available to an individual and his propensity for crime is examined in Section V. Section VI examines inter-group differences in crime rates. This section also shows why crime might spill over across groups, and why desegregation may affect different groups' crime rates. Section VII presents some extensions of the earlier analysis, as well as some concluding remarks.

II. A MODEL OF INDIVIDUAL PERCEPTION AND CHOICE

This section first presents the model, and then derives some properties of an individual's perception and choice. Several simplifying assumptions that are employed here are later relaxed.

Consider an individual who begins his active life in period t. The individual's characteristics, to be described later, are denoted by the vector h. Let $p(t, T, h)$ denote his estimate, at the beginning of period T, of the probability of punishment if he chooses to be a criminal in that period. The individual is active for L periods, where $L \geq 2$.

Pay-offs. The individual's expected utility in a period is u_0 if he chooses not to be a criminal. Otherwise, it is u_2 if he is punished, or it is u_1 if he is not punished, where $u_1 > u_2$. Thus, his optimal choice is to be a criminal in period T if and only if[7]

(1) $$u \geq p(t, T, h), \text{ where } u = (u_1 - u_0)/(u_1 - u_2).$$

A natural interpretation of u is as the "relative pay-off from crime," because it is smaller if an individual has a larger pay-off from non-crime activities (that is, if u_0 is larger, or if u_1 or u_2 is smaller), or if the penalty is more severe (that is, if u_2 is smaller). It is convenient at present to treat u as an exogenous parameter that is the same for all individuals. However, as is shown later in Section VIIC, it is easy to deal with differences in u's within and across cohorts, and also with the possible endogeneity of u's, such as the possibility that a higher economy-wide crime rate might lower u. Note from (1) that if $u \geq 1$, then the individual always chooses to be a criminal, regardless of his perceptions. If $u < 0$, then the individual never chooses to be a criminal. If $1 > u > 0$, then the individual's choice depends on his perceptions. In general, there can be individuals in each of the above three categories; our analysis goes through under the reasonable assumption that part of the population is in the third category. At present, since we are using the simplification that the u's are the same for all individuals, we assume that $u_1 > u_0 > u_2$; that is, whether an individual is ex-post worse-off or better-off being a criminal depends on whether or not he is punished. Using (1), this implies $1 > u > 0$. Obviously, this assumption need not hold for all individuals in a more general model in which they may have different u's.

Definition of the Economy-wide Crime Rate. Let $c(t, T, h, u)$ denote the probability that the individual described above will choose to be a criminal in period T. We will refer to c as this individual's "propensity for crime" in period T. From (1),

(2) $$c(t, T, h, u) = \text{Prob}[u \geq p(t, T, h)].$$

We assume at present that all individuals have similar characteristics, represented by the vector h; differences in characteristics are incorporated in Section VIIA. Thus, (2) also represents the fraction of individuals in the cohort that became active in period t who will choose to be criminals in period T. Averaging (2) across cohorts, the fraction of the population that will be criminal in period T is

---

[7]In (1), an individual who is indifferent between the choices chooses to be a criminal; the alternative assumption does not alter the analysis. The individual's uncertainties concerning pay-offs are subsumed in the expected utilities u_0, u_1 and u_2. Also, it is assumed throughout the paper that the individual's choice is not strategic. This is because the economy consists of a large number of persons, and, therefore, any one individual's choice has a negligible impact on the criminal apprehension system.

(3) $$C(T) \equiv \frac{1}{L} \Sigma_{t=T-L+1}^{T} \, c(t, T, h, u) \, .$$

Thus, $C(T)$ represents the economy-wide crime rate in period T. It is assumed in (3) that each cohort has the same large number of individuals; population changes can, however, be easily included.

The Nature of Information Available to an Individual. In the introductory section, we noted the following key features of this information: that the information is limited, that a substantial part of the relevant information comes from the individual's vicinity, and that the information stochastically reflects some past values of r. Consider the following simple specification which captures these features, and which allows itself to be easily modified. In each period, the individual collects observations on n persons, where n is a positive but small fixed number. The individual observes how many of these n persons have chosen to be criminals, and how many of them, in turn, have been punished. His current perception (that is, the value of his current p), then, is influenced by the observations accumulated to date.

Formally, let the random variable $x(\tau)$ denote the number of criminals an individual has observed, out of n persons, in period τ. Let the random variable $y(\tau)$ denote the number of criminals, out of $x(\tau)$, who were punished. For notational brevity, define vectors $\bar{x}(t, T) \equiv (x(t), ..., x(T-1))$ and $\bar{y}(t, T) \equiv (y(t), ..., y(T-1))$. Thus, an individual's perception, $p(t, T, h)$, is influenced by the value of the random vectors $\bar{x}(t, T)$ and $\bar{y}(t, T)$ he has observed.[8] Therefore, the individual's perception can be described by a reduced-form function, P, of $\bar{x}$ and $\bar{y}$.[9] That is,

(4) $$p(t, T, h) = P(\bar{x}(t, T), \bar{y}(t, T), h) \, .$$

As we shall see, Bayesian inference is a special case of this formulation of the individual's perception. Using (2) and (4), then, an individual's propensity for crime in period T can be expressed as

(5) $$c(t, T, h, u) \equiv \mathrm{Prob}[u \geq P(\bar{x}(t, T), \bar{y}(t, T), h)] \, .$$

The above specification of the information available to an individual can be modified in several ways.

---

[8] For simplicity, it is assumed here that a criminal's punishment plays the same informational role no matter who that criminal is. One possible modification is that an individual may give a different weight to his personal experience of punishment than to observations concerning others. Another simplification used here is that a punished criminal returns to the population in the next period. This can be modified by including incarceration. Also, for such reasons as the possible effects of his current choice on his future gains from crime, an individual's choice, (1), may need to be described by a dynamic program. However, these and other modifications discussed later do not alter the key qualitative properties of individual behavior, such as (6), to be described below.

[9] To confirm that plausible and consistent models of individual learning will yield an estimate such as (4), consider the following model. An individual assumes that, within different periods of his concern, the actual probability of punishment, r, is approximated by an unknown fixed number. In the beginning of his active life, his beliefs about this number are represented by a non-degenerate prior probability distribution, and he updates his beliefs in each period based on the information available up to that time. Therefore, given the nature of the variable about which the individual is attempting to learn, his stochastic observations are consistent with his initial assumption that r is approximated by an unknown fixed number.

The mass media can be modeled by defining another category of data which, though common to all individuals, contains virtually no relevant information for inferring r. Hearsay or indirect observations can be modeled by defining yet another category of data which contain potentially serious errors of observation and communication. In fact, as we noted earlier, no observation is error-free, except one's own past experience, because others' experiences not only reflect the past values of r but also others' characteristics (for instance, their ability to escape punishment). Insofar as an individual's inference of r is concerned, the net effect of a given number of observations containing errors is nearly the same as that of a smaller number of error-free observations. This aspect is captured in the specification in (4) in which the observations are depicted as error-free.

Furthermore, the number of persons from whom an individual collects observations in each period, n, may be a random number. Or, this number may be partly chosen by the individual, given relevant costs and benefits. In this case, an appropriate assumption to make, in the light of the earlier discussion, would be that while the cost is low for a small number of observations, it increases rapidly as n increases. Also, the nature of the observations may be partly state-dependent. For example, if a punished criminal is incarcerated, then his knowledge may be influenced by his closer, but not always accurate, access to other criminals' experiences.

However, the main point is not one or the other specification per se. What is crucial is to focus on the key features of the nature of the information available to an individual. Our analysis draws upon these features. The specification described above is merely a simple way to capture these features. Also, this specification points to the ways in which more elaborate and realistic models of an individual's inference can be constructed.

' Two related points may be noted here. First, it is apparent from the specification in (4) that an individual who has been active for a larger number of periods will have more information concerning the past values of r. If the value of r is not changing much over time, then this individual's current estimate, p, can be quite close to the current value of r. By the same token, there can be considerable variance in the perceptions of those in younger cohorts. Our model does not put any restriction on the length of an individual's active life-span, L. In reality, however, one would not expect L to be very large. We do not deal with the case in which L is infinite; this is realistic in itself; also, it allows us to avoid the purely technical problem of dealing with difference equations with infinite lags.

Second, we can easily modify the specification in (4) to analyze the perceptions of individuals in

different societal groups. This is done in Section VI, where an individual collects, in each period, observations from a different number of persons in different groups, and where the number of observations from one's own group may be larger. This description captures a key feature noted earlier, namely, that a substantial part of the relevant information comes to an individual from his own vicinity.

A Property of the Individual's Perception. We assume a mild form of rationality in an individual's perception, that his current p is higher if a larger proportion of those criminals whom he has observed in any past period has been punished. That is,

(6) $$\partial P/\partial x(\tau) < 0 \quad \text{and} \quad \partial P/\partial y(\tau) > 0 \quad \text{for} \quad T-1 \geq \tau \geq t .$$

As we shall see, this property is automatically satisfied if the individual's perception is described by Bayesian inference. We now examine some of the determinants of the individual's observations, especially, past crime rates and past values of r. This, in turn, will help us analyze the properties of the individual's propensity for crime.

An individual is more likely to observe a larger number of criminals in any given period if the crime rate is higher in that period. This is because a higher crime rate in any period induces a first-order stochastic improvement in the probabilities of observing different numbers of criminals. That is,

(7) $$\frac{\partial}{\partial C(\tau)} \text{Prob}[x(\tau) \geq s] > 0 \quad \text{for any s such that} \quad n \geq s \geq 1 .$$

The above expression, as well as (9) below, is derived in the Appendix.

Actual Probability of Punishment. Let r(T) denote this probability in period T ; that is, r(T) is the fraction of criminals actually punished in that period. Since r(T) is an output of the criminal apprehension system, it can be described by a transformation function. The resources spent on the criminal apprehension system in period T, denoted by E(T), is an input to this transformation function, whereas the crime rate in that period is a "negative input." Thus,

(8) $$r(T) = r(C(T), E(T)) , \quad \text{where} \quad r_C(T) < 0 \quad \text{and} \quad r_E(T) > 0 .$$

That is, r(T) is smaller if more crimes have occurred in period T or if fewer resources have been spent on the criminal apprehension system. This is because, in either case, fewer resources are available for apprehending each criminal. In (8), and elsewhere in the paper, a letter subscript represents the variable with respect to which a partial derivative is being taken. It is assumed throughout that $1 > r(T) > 0$.

The effect of r on an individual's observations is easily ascertained. If r(T) is larger, then an indi-

vidual is more likely to find a larger number of criminals punished within the subset he has observed in period T. This is because a larger $r(T)$ induces a first-order stochastic improvement in the probabilities of finding different numbers of criminals being punished. That is,

(9) $\qquad \dfrac{\partial}{\partial r(\tau)}$ Prob$[y(\tau) \geq z] > 0$ for any z such that $x(\tau) \geq z \geq 1$.

Properties of an Individual's Propensity for Crime. Expression (7) shows how the crime rate $C(\tau)$ influences the probabilities of different values of $x(\tau)$. Expression (9) shows how $r(\tau)$ influences the probabilities of different values of $y(\tau)$. From (5), therefore, an individual's propensity for crime can be expressed as a reduced-form function, g, of the crime rates and the values of r in the relevant past periods. That is,

(10) $\qquad c(t, T, h, u) = g(C(t), ..., C(T-1), r(t), ..., r(T-1), h, u)$,

where, in the right-hand side, the C's and r's play a role in determining the individual's current perception.

We now evaluate how the individual's current propensity for crime is affected by past crime rates, past values of E, and the current pay-off from crime. The following properties of (10), derived in the Appendix, play a significant role in this analysis:

(11) $\qquad \partial g/\partial r(\tau) < 0$ and $\partial g/\partial u > 0$.

That is, an individual's propensity for crime is higher if r was lower during a past period of his active life, or if the current relative pay-off from crime is higher. The reason for the latter effect is straightforward. The former effect arises because a lower r makes it more likely that a smaller number of criminals are punished within any subset of criminals. In turn, an individual will be more likely to lower his current and future p's. Consequently, his propensity for crime will increase.

To examine the effect of E's, we use (8), (10) and (11) to obtain

(12) $\qquad dg/dE(\tau) = r_E(\tau)\partial g/\partial r(\tau) < 0$.

That is, an individual has a higher current propensity for crime if fewer resources were spent on the criminal apprehension system during a past period of his active life. The intuition behind this result is simple. Fewer resources spent on the criminal apprehension system dilute the resources spent on apprehending each criminal. The resulting decline in r leads, for the reasons described in the previous paragraph, to an increase in the individual's propensity for crime.

Next, consider the effect of past crime rates. Expressions (8) and (10) yield

$$(13) \qquad dg/dC(\tau) = r_c(\tau)\partial g/\partial r(\tau) + \partial g/\partial C(\tau) .$$

The two terms in the right-hand side describe two different effects that a higher past crime rate has on an individual's propensity for crime. The first term keeps unchanged the number of criminals that the individual has observed, and evaluates the effect that a higher crime rate has through its impact on $r(\tau)$. The second term keeps $r(\tau)$ unchanged, and evaluates the impact that a higher crime rate, $C(\tau)$, has through its impact on how many criminals an individual is likely to observe.

The first term is unambiguously positive from (8) and (11). The explanation is as follows. A higher crime rate in a period dilutes the resources spent on apprehending each criminal. This leads to a decrease in r in that period. Consequently, for reasons described earlier, an individual's propensity for crime increases. The second term in the right-hand side of (13), namely,

$$(14) \qquad \partial g/\partial C(\tau) ,$$

is evaluated in the Appendix. It is shown there that though (14) can in general be positive or negative, it is non-negative for small values of $r(\tau)$. Now, in practice, the value of r is very small. For instance, in the U.S., it is estimated that fewer than 33 percent of crimes are reported. Of them, fewer than 20 percent lead to an arrest. Further, even for serious crimes, significantly fewer than half of the arrests lead to punishment (see Bureau of Justice Statistics (1983)). Such estimates, although not robust for reasons noted in the introductory section, indicate that, for most crimes, the actual probability of punishment has a very small order of magnitude.

Now, recall that the first term in the right-hand side of (13) is unambiguously positive. Thus, given the arguments presented in the previous paragraph, it is reasonable to conclude that (13) is positive in practice. That is,

$$(15) \qquad dg/dC(\tau) > 0 .$$

The conclusions obtained in this subsection are summarized in

PROPOSITION 1. *An individual's current propensity for crime is higher, if during a past period of his active life (i) the crime rate was higher, or (ii) fewer resources were spent on the criminal apprehension system, or if (iii) the current relative pay-off from crime is larger.*

III. THE ECONOMY-WIDE CRIME RATE

This section derives the current economy-wide crime rate as a function of some past and present variables. It then identifies some of the properties of the crime rate and examines the impact of a change in the economy's parameters on the crime rate.

Aggregation of (10) across cohorts, according to (3), allows us to define a reduced-form function, f, described below, for the current economy-wide crime rate:

(16) $$C(T) = \frac{1}{L}\Sigma_{t=T-L+1}^{T} c(t, T, h, u)$$
$$= f(C(T-1), ..., C(T-L+1), E(T-1), ..., E(T-L+1), h, u, \text{other parameters}) .$$

Using (10), (11), (12) and (15), we can obtain the following derivatives of (16):

(17) $$\partial C(T)/\partial C(T - \tau) > 0 \text{ for } L - 1 \geq \tau \geq 1 ,$$

(18) $$\partial C(T)/\partial E(T - \tau) < 0 \text{ for } L - 1 \geq \tau \geq 1 , \text{ and}$$

(19) $$\partial C(T)/\partial u > 0 .$$

Expressions (17) and (18) show how the current crime rate is affected by the crime rates and the resources spent on the criminal apprehension system during the past $L - 1$ periods. The number of periods matters here only because the oldest cohort currently active has been active for $L - 1$ periods. However, it is apparent from (16) that, as in any dynamic system, the variables from a more distant past period can have indirect effects on the current outcome, even though the magnitude of the effects may decline over time. For example, suppose the crime rate was higher in a past period. Then, through (16), the crime rates in the immediate subsequent periods will be higher, and this chain of influence will be felt on a future crime rate, even though the magnitude of the influence may keep declining. Using (17), (18), and (19), thus, we obtain the following economy-wide analogue of Proposition 1.

PROPOSITION 2. *The current crime rate is higher, if in a past period, (i) the crime rate was higher, or (ii) fewer resources were spent on the criminal apprehension system, or if (iii) the current relative pay-off from crime is larger.*

Stable Steady-State Crime Rates. As discussed in the introductory section, a useful device for examining some aspects of the dynamic relationship in (16) is to look at its steady-states. We focus on the stable steady-states because if the crime rate is close to a stable level, then subsequent to small shocks,

future crime rates will not diverge away rapidly. Throughout this paper, stability refers to local stability (see Hirsch and Smale (1974, pp. 185-87, 278-81), Sah (1989a), and Whittle (1983, pp. 46-50) for standard definitions). Also, the steady-states on which we focus are interior; that is, the crime rate, C(T), at these steady-states is larger than zero but smaller than one.

Note that, given the feedback property identified above, that past crime breeds future crime, one might inadvertently conclude that the economy must eventually gravitate towards complete criminality, because any shock which raises the crime rate will induce higher and higher future crime rates. If this were the case, then stable interior steady-states would be ruled out. To see intuitively why this is not the case, suppose the economy is currently at an interior steady-state with crime rate C, and that a one-time exogenous shock has raised the current crime rate to $C + \Delta C$. Now, recall that a new cohort enters the economy in each period, and the oldest cohort leaves. Hence, the fraction of the population which has become criminal due to the shock will leave the economy at some future time. Consider a period after this. We know from Proposition 2(i) that the crime rate in this period will be higher than C, but there is no restriction on how much higher it will be. In particular, it can be smaller than $C + \Delta C$. In such cases, the economy will gravitate towards the original crime rate C. Thus, stable interior steady-states are unambiguously possible.

The reason for this conclusion is simple. The feedback property in (17) does not rule out stable interior steady-states or, for that matter, unstable interior steady-states. Moreover, as is shown later in Section VIIB, it is easy to identify mild conditions that guarantee one or more stable interior steady-states.

Some Properties of the Crime Rate. Let C denote a stable interior steady-state crime rate. Let E denote the unchanging level of resources spent per period on the criminal apprehension system. In (16), substitute C for $C(\tau)$'s, and E for $E(\tau)$'s. Let F denote the resulting value of f. Note, for later use, that F represents the current crime rate, as determined by past variables. In a steady-state, F must equal C. Thus, steady-state crime rates are those values of C that satisfy the equation

(20) $C = F(C, E, h, u, \text{other parameters})$.

In general, the function F is highly non-linear in C, as we shall see in the next section. Equation (20) will thus admit multiple values of C as solutions. A consequence of this is

PROPOSITION 3. *Two economies with nearly identical current parameters can have different steady-state crime rates.*

This result follows from the earlier dynamic analysis. If two economies have identical parameters in the current and future periods, but one of them has had higher crime rates in some recent past periods, then this economy will have higher future crime rates.

For comparative statics of a steady-state, let θ denote a parameter, and let (20) be stated as: $C = F(C, \theta)$. Perturbation of this equation with respect to a sustained change in θ yields $dC/d\theta = F_\theta/(1 - F_C)$. These comparative statics are in the neighborhood of a stable steady-state. Note that (17) implies that $F_C(C) > 0$. Also, C satisfies $1 > F_C(C)$, which is a necessary condition for C to be stable (see Sah (1989a)). We thus obtain

(21) $$\text{sgn}\{dC/d\theta\} = \text{sgn}(F_\theta) \quad \text{and} \quad |dC/d\theta| > |F_\theta|.$$

For an interpretation of these results, recall that F is the current crime rate as determined by past variables. F_θ can thus be viewed as the "first-round" or the immediate impact of a change in θ on the crime rate. Future crime rates are altered not only by the changed value of θ, but also by a sequence of indirect effects. The difference between the post- and pre-change steady-state crime rates is described by $dC/d\theta$. Expression (21) thus implies the following relationship between the overall change in the crime rate and the first-round impact of a parameter change:

PROPOSITION 4. *The first-round impact of a parameter change on the crime rate is preserved in sign but amplified in magnitude as a new steady-state crime rate is approached.*

This result holds no matter which parameter changes. As an example, suppose that there is a sustained increase in the relative pay-off from crime, u. Then, from (16), (19) and (20), $F_u > 0$. That is, the first-round impact of this parameter change is an increase in the crime rate, because the increased attractiveness of crime will induce some individuals to alter their choice. This, in turn, will make it more likely that individuals in the future will have lower p's. This will increase their propensity for crime. The indirect effect thus reinforces the first-round impact.

IV. BAYESIAN INFERENCE AS A DESCRIPTION OF INDIVIDUALS' PERCEPTIONS

The preceding sections employed a relatively general formulation to describe individuals' perceptions. One description that is consistent with this formulation is Bayesian inference. This section analyzes crime rates under this specification. The results are useful later as well.

Consider the perception, in period T, of an individual who became active in period t. Since he has been active for $\ell = T - t$ periods, he has collected observations on ℓn persons. Among these ℓn persons, let $X(t, T) = \Sigma_{\tau=t}^{T-1} x(\tau)$ denote the number of criminals. Among these, let $Y(t, T) = \Sigma_{\tau=t}^{T-1} y(\tau)$ denote the number of those who are punished. Since Y is a sufficient statistic given any X, Bayesian inference implies that the individual's estimate, $p(t, T, h)$, is influenced by $X(t, T)$ and $Y(t, T)$, but not by how these two random numbers are distributed in different periods. Consequently, expression (4) simplifies to

(22) $p(t, T, h) = P(X(t, T), Y(t, T), h)$.

Using standard Bayesian techniques, it is easily established that

(23) $P_X < 0$, and $P_Y > 0$.

That is, an individual has a lower p if a smaller overall proportion of criminals, among those he has observed, has been punished. Thus, condition (6) is automatically satisfied by (22).

For an intuitive description of the individual choice, define a function Z through the equality:

(24) $u = P(X, Z(X, h, u), h)$.

Then from (1), (22), (23) and (24), it follows that an individual will choose to be a criminal if and only if:

(25) $Y \le Z(X, h, u)$.

That is, an individual will be a criminal if and only if he has found no more than Z criminals punished among the X criminals whom he has observed.

A natural interpretation of Z is as a "reservation level." From (23) and (24),

(26) $Z_X > 0$, and $Z_u > 0$.

That is: If an individual has observed a larger number of criminals, or if the relative pay-off from crime is larger, then he may observe a larger number of criminals being punished and still choose to be a criminal.

An individual's propensity for crime can be described as follows. Since C is the crime rate, the probability that the individual has observed X criminals is

(27) $b(X, \ell n, C) = \binom{\ell n}{X} C^X (1 - C)^{\ell n - X}$,

which is the binomial probability of X successes out of ℓn trials, where C is the probability of success in each trial.

The actual probability of punishment is now

(28) $\qquad r = r(C, E)$, where $r_C < 0$ and $r_E > 0$.

Therefore, the probability that an individual will observe Z or fewer criminals punished, out of X, is

(29) $\qquad B(Z, X, r) = \sum_{j=0}^{Z} \binom{X}{j} r^j (1 - r)^{X-j}$,

which is the binomial cumulative probability of Z or fewer successes out of X trials, where r is the probability of success in each trial. Thus, recalling (25), it is clear that (29) is the probability that an individual who has observed X criminals will choose to be a criminal. Note that (29) is defined for an integer Z, whereas (24) is not. However, we suppress this qualification to keep the exposition uncluttered. As can be confirmed from the second and third parts of the Appendix, this does not affect the qualitative analysis. Also, for brevity, we have suppressed the arguments of the function Z in (29) and in the rest of the paper. That is,

(30) $\qquad Z = Z(X, h, u)$.

Combining (27) with (29), and summing over X, the probability that an individual who became active in period t will be a criminal in period T is

(31) $\qquad c(t, T, h, u) = \sum_{X=0}^{\ell n} b(X, \ell n, C) B(Z, X, r)$,

where $\ell = T - t$. This describes the individual's propensity for crime.

Aggregation of (31) across cohorts, according to (3), yields the desired equation for steady-state crime rates:

(32) $\qquad C = \frac{1}{L} \sum_{\ell=0}^{L-1} \sum_{X=0}^{\ell n} b(X, \ell n, C) B(Z, X, r)$,

where expressions (28) and (30) respectively describe r and Z. Clearly, (32) is an explicit version of the more general expression in (20).

A Special Case. One of the determinants of an individual's current perception is his set of initial beliefs; that is, the beliefs concerning the value of r with which he begins his active life. The analysis presented above holds for all types of non-degenerate initial beliefs. Suppose now that initial beliefs are represented by a beta distribution. This specification, which is sometimes used below, is not particularly restrictive, because most types of initial beliefs can be closely approximated by a suitably chosen beta distribution. Let the elements h_1 and h_2 of the vector of characteristics h denote the parameters of the

above beta distribution. By definition, h_1 and h_2 are positive. Then, from a standard result on beta distribution (see DeGroot (1970, p. 160)), (22) can be stated as

(33) $\qquad p = (h_1 + Y)/(h_1 + h_2 + X) .$

This and (24) yield a closed-form expression for the reservation level Z :

(34) $\qquad Z = (h_1 + h_2 + X)u - h_1 .$

Note from (33) that an individual's initial estimate of the probability of punishment is $h_1/(h_1 + h_2)$, which is obtained by setting $Y = X = 0$. Further, recalling that $1 > u > 0$, (34) yields $\partial Z/\partial h_1 < 0$, and $\partial Z/\partial h_2 > 0$. Thus, an implication of (34), which is also used later, is that a smaller initial p , due to a smaller h_1 or a larger h_2 , raises future reservation levels. That is: If an individual initially believes the probability of punishment to be smaller, then in the future he might observe a larger maximum number of actual punishments and still choose to be a criminal.

V. THE EFFECTS OF AN INDIVIDUAL'S OBSERVATIONS ON HIS CHOICE

Throughout this paper, we have emphasized the role the information available to an individual plays as a determinant of his perception and choice. It is useful to examine how an individual's choice might change if the nature of his information changed. An example of such an analysis is presented in this section. This analysis is also useful later when an economy consisting of dissimilar groups is examined.

Suppose an individual observes $n + 1$ persons, rather than n , in each period. This is a simple way to examine what happens if the individual has "more" information. Since the individual observes more persons, the probability that he will observe more criminals is larger. This is because a larger n induces a first-order stochastic improvement in the probability density (27). Does this change raise or lower his propensity for crime? The following three considerations are important in determining the answer.

(i) How large is r ? A smaller r means that fewer of the additional criminals observed by the individual are likely to be punished. This is likely to reduce his p and thereby increase his propensity for crime. (ii) How large is the relative pay-off from crime, u ? For any given change in his observations, the individual is more likely to choose to be a criminal if u is larger. (iii) What is the nature of the individual's initial beliefs? An individual with stronger initial beliefs is less likely to alter his choices due to changed observations.

A simple framework to assess the outcome of these three considerations together is one in which the individual uses the Bayesian inference described in (33), and is initially indifferent about whether to be a criminal or not. Then, as shown in the Appendix, we obtain

PROPOSITION 5. *A greater number of observations raises (lowers) an individual's propensity for crime, if the relative pay-off from crime is larger (smaller) than the actual probability of punishment.*

VI. CRIME RATES IN DIFFERENT SOCIETAL GROUPS

This section examines an economy with two societal groups. The analysis presented here can be extended to deal with many groups. As a part of this analysis, we examine how crime might spill over from one group to another, and how a change in the degree of inter-group segregation might affect different groups' crime rates. Superscripts $i = 1$ and 2 denote the variables for the two groups.

Recall from the introductory section that a key feature of the information available to an individual is that a substantial part of the relevant information comes to him from his immediate vicinity. A natural way to describe this feature within our specification is that, out of the n^i observations that an individual in group i collects in each period, a significant fraction come from his own group. Let m^i denote the number of observations that a group-i individual collects from the first group. That is, in each period, an individual belonging to the first group collects m^1 observations from his own group, whereas an individual belonging to the second group collects $n^2 - m^2$ observations from his own group. Thus, a larger value of m^1 or a smaller value of m^2 means that one's own group is a more important source of information to an individual. If $m^1 = n^1$ and $m^2 = 0$, then the two groups are completely segregated, so far as the information flow to individuals is concerned. The individuals in the two groups may also differ in their characteristics, which are denoted by the vector h^i for a group-i individual. The relative pay-off from crime to a group-i individual is denoted as u^i. We continue to abstract from intra-group differences in characteristics and pay-offs, but these are easily incorporated as discussed in Section VII.

Let C^i denotes the crime rate in group i in a stable steady-state. Then, analogous to (20), we can express the steady-state equation system as:

$$(35) \qquad C^1 = F^1(C^1, C^2, \theta^1), \text{ and } C^2 = F^2(C^1, C^2, \theta^2),$$

where θ^i denotes a parameter affecting individuals in group i. Among the parameters are individuals' characteristics and pay-offs, and the parameters n^i and m^i describing the nature of their observations.

Bayesian Inference. If individuals' perceptions are described by Bayesian inference, then, by going through the following steps, we can obtain an explicit version of (35). First, if α denotes the fraction of the population belonging to the first group, and if the actual probability of punishment is the same for both groups, then, instead of (28), r is given by

$$(36) \qquad r = r(\alpha C^1 + (1 - \alpha)C^2, E) .$$

Note that (36) is easily generalized, for instance, by defining $r^i = r^i(C^1, C^2, E)$ to be the group-specific actual probability of punishment. Second, instead of (30), the reservation level for an individual in group i is now $Z^i = Z(X, h^i, u^i)$. This reflects the differences between individuals' characteristics and pay-offs in the two groups. If the initial beliefs of individuals are described by a beta distribution, then, analogous to (34), we obtain

$$(37) \qquad Z^i = (h_1^i + h_2^i + X)u^i - h_1^i .$$

Third, we need the counterpart of (27) to describe the probability that a group-i individual who has been active for ℓ periods has observed X criminals. Since a group-i individual collects m^i observations from the first group and $n^i - m^i$ observations from the second group, this probability is

$$(38) \qquad \phi^i(X, \ell) = \Sigma_{j=0}^{\ell m^i} b(j, \ell m^i, C^1)b(X - j, \ell n^i - \ell m^i, C^2) .$$

Using these components, we can derive, analogous to (32), the following steady-state equation system:

$$(39) \qquad C^i = \frac{1}{L}\Sigma_{\ell=0}^{L-1} \Sigma_{X=0}^{\ell n^i} \phi^i(X, \ell)B(Z^i, X, r) , \text{ for } i = 1 \text{ and } 2.$$

Complete Segregation. Recall that, for the purpose at hand, a higher degree of segregation means that the perceptions of individuals in one group are less affected by the behavior of individuals in the other group. In the extreme case of complete segregation between groups, the behavior and dynamics of one group does not affect the other. Since Proposition 3 applies to each group in such a case, we obtain

PROPOSITION 6. *Highly segregated groups can have different crime rates, even if members of all groups face identical parameters. Therefore, significant differences between the crime rates of different groups may coexist with relatively modest differences between their economic parameters.*

Spillover of Crime across Groups. Now consider those situations in which there is some inter-group interaction. A consequence of inter-group interactions is that the perceptions of the members of one group are influenced, to some degree, by what individuals in the other group do. Thus, one group's crime rate can have spillover effects on other groups. We examine these effects as follows. Suppose there is a change in a parameter faced by the members of one group, but no change in the parameters faced by the members of another group. Then, what is the impact on the crime rate in the latter group?

Let the first-round impact of an increase in past crime rates, $\partial F^i/\partial C^j$, be denoted as F^i_j, where $i = 1$ and 2, and $j = 1$ and 2. Using reasoning similar to that which yielded (17), we obtain $F^i_j > 0$. Then, two implications of (35), derived in the Appendix, are as follows. First,

(40) $\qquad \text{sgn}\{dC^1/d\theta^i\} = \text{sgn}\{dC^2/d\theta^i\}$.

Second, if the parameter change is in the neighborhood of the case in which the two groups face the same set of parameters, then

(41) $\qquad |dC^i/d\theta^i| > |dC^j/d\theta^i|$, for $i \neq j$.

These results are respectively summarized in

PROPOSITION 7.

> *(i)* *If a change in parameters faced by one group's members raises this group's crime rate, then the crime rate in the other group also rises because of spillovers, even though the latter group's members have not experienced any change in parameters.*

> *(ii)* *If differences between the parameters faced by two groups are small, then the change in the crime rate of the group whose members face a parameter change exceeds in magnitude the spillover impact on the other group's crime rate.*

These results hold for all types of parameter changes. For example, suppose the relative pay-off from crime has increased for the members of the first group. Then, it is apparent that this group's crime rate will rise. In addition, because of inter-group interactions, and the resulting change in individuals' perceptions and in the actual probability of punishment, the choices of those in the second group will be altered such that this group's crime rate will rise as well. Moreover, if differences in the parameters faced by the members of the two groups are small, then the first group will experience a larger crime rate increase than the second group.

The Effects of a Change in Segregation. To evaluate these effects, define a variable S such that a larger S implies greater segregation. That is,

$$(42) \qquad m^1 = M^1 + S, \text{ and } m^2 = M^2 - S,$$

where M^1 and M^2 are pre-specified numbers. Suppose that the only parameter which differs between the members of the two groups, other than differences in the nature of their observations, is the relative pay-offs from crime. Let $u^1 > u^2$. Such differences in pay-offs can arise for many reasons. For brevity, however, we shall use the verbal interpretation that members of the first group are poorer; that is, they have less attractive non-crime opportunities. Let the poorer group's crime rate exceed that of the richer group; this will be the case if u^1 is sufficiently larger than u^2.

Now, consider a decrease in the degree of segregation resulting from a unit decrease in S. That is, a poorer person now collects one more observation from the richer group and one fewer from the poorer group. The opposite is the case for a richer person. Then, as is shown in the Appendix:

(43) Decreased segregation will make it more likely that a poorer person will observe

fewer criminals and that a richer person will observe more criminals.

The impact of this change on individuals' choices is determined by the three considerations described in Section V. In addition, we need to take account of the spillovers of the type identified in Proposition 7(i). For instance, if the poorer group's crime rate falls because of the change in the perceptions of *its* members, then the spillover will lower the richer group's crime rate. Likewise, if the richer group's crime rate rises because of the change in the perception of *its* members, then the spillover will raise the poorer group's crime rate.

Now, consider the case in which individuals' perceptions are described by the Bayesian inference in (37), and they are initially indifferent about whether to be a criminal or not. Then, as shown in the Appendix, we obtain

PROPOSITION 8. *If the relative pay-off from crime for those in the poorer group exceeds the actual probability of punishment, and the latter exceeds the relative pay-off from crime for those in the richer group, then a decrease in the degree of segregation lowers both groups' crime rates.*

The intuition behind this result is as follows. Recall the effects of a smaller S, described in (43). Under the conditions of Proposition 8, the first-round impact of these effects decreases the propensity that

individuals in either group will have for crime. This decrease is reinforced by indirect effects within a group and by the spillovers across groups, as described respectively by Propositions 4 and 7(i). The overall consequence is a decrease in both groups' crime rates.

VII. EXTENSIONS AND REMARKS

To keep the exposition uncluttered, a number of simplifying assumptions were made in the earlier sections. This section relaxes some of these assumptions and presents some remarks. These extensions also suggest some additional insights.

A. Differences in Individuals' Characteristics

We have thus far abstracted from differences in individuals' characteristics within a group. For a more general model, let $dK(h)$ denote the fraction of each cohort having characteristics represented by the vector h. Then the definition of the crime rate for a single group economy, (3), becomes

$$(44) \qquad C(T) = \frac{1}{L} \Sigma_{t=T-L+1}^{T} \int c(t, T, h, u) dK(h) \ .$$

This modification does not alter the reduced-form expressions, (16) and (20), for the crime rate, but alters equation (32) to

$$(45) \qquad C = \frac{1}{L} \Sigma_{\ell=0}^{L-1} \Sigma_{X=0}^{\ell n} \int b(X, \ell n, C) B(Z, X, r) dK(h) \ .$$

Analogous extensions apply to economies with several groups.

An illustration of the impact of a change in individuals' characteristics is as follows. Consider a change in the distribution of initial beliefs. Recall from the discussion relating to (34) that a smaller h_1 means that the individual has a lower initial p. Now, suppose that due to a first-order stochastic worsening in the distribution of h_1, larger fractions of the population now have lower initial p's. It can be shown that such a change raises the crime rate. Correspondingly, from (28), r is lower. Thus, we obtain

PROPOSITION 9. *If more individuals initially have smaller perceived probabilities of punishment, then not only does the subsequent crime rate rise, but also the actual probability of punishment decreases.*

B. Remark on Stable Steady-States

In Section III, we briefly discussed why the dynamic systems developed in this paper, such as (20), admit stable interior steady-states. Here, we illustrate how one can identify mild conditions that are sufficient to guarantee such steady-states.

Suppose individuals are active for two periods; that is, $L = 2$. Then, the relationship in (16) between $C(T)$ and $C(T - 1)$ can be depicted as in Figure 1. The necessary and sufficient condition for stability in this case is: $1 > \partial C(T)/\partial C(T - 1)$; that is, the function f in Figure 1 should intersect the 45° line from above. The following result is then easily established: A sufficient condition for one or more stable interior steady-states is that at least one person in each cohort begin his active life as a criminal, and at least one as a non-criminal.

To prove this result, recall the formulation underlying (44). If at least one person begins his active life as a criminal, and at least one as a non-criminal, then a cohort's crime rate in its first active period is larger than zero but smaller than one. That is: $1 > \int c(T, T, h, u)dK(h) > 0$. Now, substitute this and $L = 2$ into (44), and note that $1 \geq c(T - 1, T, h, u) \geq 0$, because c denotes a probability. This yields $1 > C(T) > 0$, for $1 \geq C(T - 1) \geq 0$.

Thus, f intersects the two vertical axes in Figure 1 at interior points such as A and A', guaranteeing at least one stable interior steady-state. The condition mentioned above for this conclusion is sufficient but not necessary. This is because even if f were to touch the two vertical axes in Figure 1 at corners O and O' respectively, it would not preclude the existence of one or more stable interior steady-states.

C. Individuals' Pay-offs from Crime

For brevity, an individual's relative pay-off from crime was treated earlier as a fixed parameter, u, which was assumed to be the same for all individuals within a group. This aspect is easily modified. For instance, u might be affected by other variables in the economy. In particular, one might expect u to become smaller at higher crime rates, because of the crowding out of the opportunities from crime. That is, $u = u(C)$, where $u_C < 0$. Analogous extensions can deal with the effects of the crime rate on other variables that have thus far been treated as parameters.

Furthermore, u's can be different for different individuals within a group, reflecting their: (i) age (because the punishment of adults is often different from that of younger criminals, for instance), (ii) abilities, income, employment and other determinants of crime versus non-crime opportunities, (iii) tastes, including the discount factors for calculating personal losses from future punishments, and (iv) past crime history, which may affect an individual's productivity in crime (through learning by doing, for instance) as well as his non-crime opportunities (due to the stigma attached to convicts).

To see an illustration of the impact of changes in u's, let the characteristic h_3 denote an individual's income, such that $u = u(h_3)$, and $\partial u(h_3)/\partial h_3 < 0$. That is, other things being equal, poorer persons
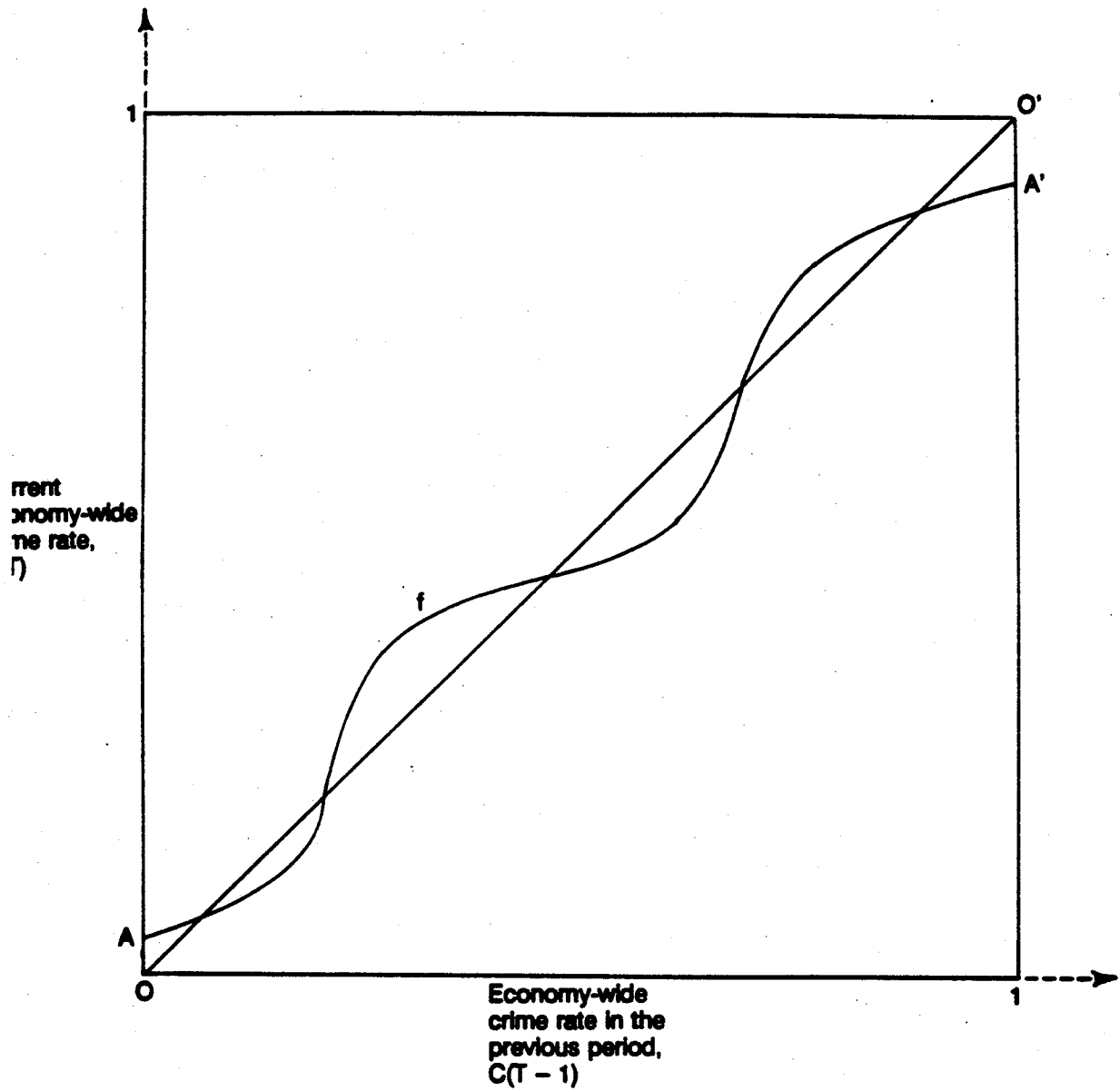
FIGURE 1

Illustration that a sufficient condition for one or more stable interior steady-states is that at least one person in each cohort begin his active life as a criminal, and at least one as a non-criminal.

find crime more attractive; the opposite case (in which, for instance, a richer criminal can better protect him-self against punishment) can be worked out just as easily. Now, consider a first-order stochastic worsening in income distribution; that is, larger fractions of the population now have smaller incomes. Then, it is easily shown that such a change in income distribution will raise the crime rate.

APPENDIX

Derivation of (7) and (9). Let j denote a binomial variate with parameters (n, C). Let $b(j, n, C) = \binom{n}{j}C^j(1 - C)^{n-j}$ denote the probability density of j. Let $B(s, n, C) = \sum_{j=0}^{s} b(j, n, C)$ denote the corresponding cumulative density. We follow the natural convention that for $1 > C > 0$:

(A1) $b(j, n, C) = 0$ if $j < 0$ or if $j > n$, $B(s, n, C) = 0$ if $s < 0$, and

$B(s, n, C) = 1$ if $s \geq n$.

Note that B satisfies the following identify (see Sah (1989b) for a general version of this identity):

(A2) $B(s, n + 1, C) - B(s, n, C) = -Cb(s, n, C)$.

Another property of B that we use later is

(A3) $\partial B(s, n, C)/\partial C < 0$ for $s = 0$ to $n - 1$.

Now, by definition, $\text{Prob}[x(\tau) \geq s] = 1 - B(s - 1, n, C(\tau))$. This and (A3) yields (7). The derivation of (9) is analogous.

Derivation of (11). For brevity, we derive (11) for $T = t + 1$; the same logic holds for other T's. Also, in this derivation and in the evaluation of (14) below, we suppress h, and denote $x(t)$ as x, $y(t)$ as y, and $r(t)$ as r. Thus, from (5), $c = \text{Prob}[u \geq P(x, y)]$. It is assumed throughout that $1 > c > 0$. Define a function $Z(x, u)$ such that $u = P(x, Z(x, u))$. Let $z(x, u)$ denote the largest integer not larger than $Z(x, u)$. Then, using (6) and the definition of z, we obtain

(A4) $z(x + 1, u) \geq z(x, u)$, and

(A5) $z(x, u') \geq z(x, u)$ for $u' > u$.

Also, (6) and the definitions of z and c yield: $c = \text{Prob}[y \leq z(x, u)]$. Thus, (10) can be expressed as

(A6) $c = \sum_{x=0}^{n} \text{Prob}[x] \cdot B(z(x, u), x, r)$.

Next, define sets I_1, I_2 and I_3 such that: $x \in I_1$ if $z(x, u) \geq x$, $x \in I_2$ if $x - 1 \geq z(x, u) \geq 0$, and $x \in I_3$ if $z(x, u) < 0$. Then, using (A1), we can restate (A6) as

(A7) $c = \sum_{x \in I_1} \text{Prob}[x] + \sum_{x \in I_2} \text{Prob}[x] \cdot B(z(x, u), x, r)$,

where B in the right-hand side is positive but smaller than one. Using (A3), the derivative of (A7) with respect to r yields the first part of (11).

To establish the second part of (11), note from (A5) that r is non-decreasing in u. Thus, from (A7) and from the definitions of sets I_1, I_2 and I_3, it is easily verified that the changes in these sets induced by a larger u cannot decrease c. We now make the reasonable assumption (which can be weakened) that the inequality in (A5) is strict for at least one x belonging to I_2. This yields the second part of (11).

Evaluation of (14). This evaluation is also conducted for $T = t + 1$. Also, we write $z(x, u)$ as $z(x)$. Using (A4), define a non-negative integer function $k(x)$, such that $z(x + 1) = z(x) + k(x)$. Define $\delta(x, r) = B(z(x) + k(x), x + 1, r) - B(z(x), x, r)$. Recall from (7) that an increase in $C(t)$ induces a first-order stochastic improvement in the probability density of x. Hence, from (A6) and from a standard result concerning first-order stochastic dominance (see Ingersoll (1987, pp. 137-9)), we obtain:

(A8) $\partial c / \partial C(t) \geq 0$ if $\delta(x, r) \geq 0$ for all x.

Next, we evaluate $\delta(x, 0)$. $B(z, x, r)$ has the following properties, which we use below:

(A9) $B(z, x, 0) = 0$ if $z < 0$, and $B(z, x, 0) = 1$ if $z \geq 0$.

Since $k(x) \geq 0$, three cases that exhaust all possible combinations of values of $z(x)$ and $z(x) + k(x)$ are: (i) $z(x) \geq 0$ and $z(x) + k(x) \geq 0$. In this case, $\delta(x, 0) = 0$. (ii) $z(x) < 0$ and $z(x) + k(x) < 0$. In this case as well, $\delta(x, 0) = 0$. (iii) $z(x) < 0$, but $z(x) + k(x) \geq 0$. Here, $\delta(x, 0) = 1$.

Thus, $\delta(x, 0) \geq 0$ for all x. Since $\delta(x, r)$ is simply a polynomial in r, it is continuous in r. Hence, as r tends to zero, $\delta(x, r)$ tends to $\delta(x, 0)$, which we have just seen is non-negative. In turn, (A8) yields the desired result.

Proof of Proposition 5. Recall that an increase in n induces a first-order stochastic improvement in the density b, which is defined in (27). Let Δc denote the change in c due to an increase in n. Then, using (30), (31) and the result on first-order stochastic dominance noted earlier, we obtain

(A10) $\Delta c \gtrless 0$ if $B_Z Z_X + B_X \gtrless 0$.

To analyze (A10), we approximate the binomial distribution function $B(Z, X, r)$ by $N((Z - rX)(Xr(1 - r))^{-1/2})$, where N is the unit normal distribution function (see Johnson and Kotz (1969, p. 53)). This yields

(A11) $B_Z > 0$ and $B_X = -B_Z(Z + rX)/2X$.

Next, recall from (33) that an individual's initial p is $h_1/(h_1 + h_2)$. Thus, from (1), an individual is initially indifferent about whether or not to be a criminal if $h_1/(h_1 + h_2) = u$. The last equation and (34) yield: $Z = Xu$, and $Z_X = u$. These two expressions combined with (A11) yield

(A12) $B_Z Z_X + B_X = B_Z(u - r)/2X$.

From (A11), $B_Z > 0$. Hence, from (A10) and (A12), we obtain the desired result: $\Delta C \gtrless 0$ if $u \gtrless r$.

Derivation of (40) and (41). Define $F_\theta^i \equiv \partial F^i/\partial \theta^i$, and $D \equiv (1 - F_1^1)(1 - F_2^2) - F_2^1 F_1^2$. Then perturbations of (35) with respect to θ's yield

(A13) $dC^1/d\theta^1 = F_\theta^1(1 - F_2^2)/D$, $dC^2/d\theta^1 = F_\theta^1 F_1^2/D$,

$dC^1/d\theta^2 = F_\theta^2 F_2^1/D$, and $dC^2/d\theta^2 = F_\theta^2(1 - F_1^1)/D$.

Given that $F_j^i > 0$, the following conditions are necessary for stability (see Sah (1989a)):

(A14) $1 > F_i^i$, and $D > 0$.

The result in (40) follows from (A13), (A14), and from $F_j^i > 0$. Next, if $F_j^1 = F_j^2$, then (A14) implies that $1 > F_1^i + F_2^i$. In turn, (41) is obtained by evaluating the derivatives in (A13) in the neighborhood of $F_j^1 = F_j^2$.

Derivation of (43). Recalling the definition of S, (42), the claim in (43) is a restatement of the result that: If $C^1 > C^2$, then a larger m^i induces a first-order stochastic improvement in the density ϕ^i, which was defined in (38). Recall that ϕ^i is the density of the sum of two independent binomial variables, with respective parameters $(\ell m^i, C^1)$ and $(\ell n^i - \ell m^i, C^2)$. Keeping the total sample size fixed, if a larger sample is drawn from the former distribution, then the cumulative density of the number of successes must shift to the right. See Sah (1989b) for more general relationships of this type.

Proof of Proposition 8. Note from (35) and (39) that F^i is represented here by the right-hand side of (39). Setting $\theta^1 = \theta^2 = -S$, and using (A13), then, our objective is to evaluate $dC^i/d\theta^1 + dC^i/d\theta^2$. Recall from the above paragraph that a larger m^i induces a first-order stochastic improvement in the density ϕ^i. Let ΔF^i denote the change in F^i due to an increase in m^i. Then, reasoning of the type used earlier in the proof of Proposition 5 yields: $\Delta F^i > 0$, if $u^i > r$. In turn, using (39), (A13) and (A14), the desired result follows.

REFERENCES

Andreano, R. and J. J. Siegfried, eds., 1980, *The Economics of Crime*, Wiley, New York.

Anttila, I. and R. Jaakkola, 1966, "Unrecorded Criminality in Finland," *Kriminologinen Tutkimuslaitos* 2, 5-22.

Becker, G. S., 1968, "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76, 33-66.

Becker, G. S., 1983, "A Theory of Competition among Pressure Groups for Political Influence," *Quarterly Journal of Economics* 98, 371-400.

Becker, G. S. and W. M. Landes, eds., 1974, *Essays in the Economics of Crime and Punishment*, Columbia University Press for the National Bureau of Economic Research, New York.

Bureau of Justice Statistics, 1988, *Report to the Nation on Crime and Justice*, Second Edition, U.S. Department of Justice, Washington, D.C.

Bureau of Justice Statistics, 1983, *Report to the Nation on Crime and Justice: The Data*, U.S. Department of Justice, Washington, D.C.

Carter, R. L. and K. Q. Hill, 1978, "Criminals' and Noncriminals' Perceptions of Urban Crime," *Criminology* 16, 353-371.

DeGroot, M. H., 1970, *Optimal Statistical Decisions*, McGraw-Hill, New York.

Government of France, *Compte General de l'administration de la justice civile et commerciale et de la justice criminelle*, published since 1825, Ministry of Justice, Paris.

Heineke, J. M., ed., 1978, *Economic Models of Criminal Behavior*, North-Holland, Amsterdam.

Hirsch, M. W. and S. Smale, 1974, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York.

Ingersoll, J. E., 1987, *Theory of Financial Decision Making*, Rowman & Littefield, Totowa.

Johnson, N. L. and S. Kotz, 1969, *Discrete Distributions*, Houghton Mifflin, Boston.

Klevorick, A. K., 1985, "Legal Theory and the Economic Analysis of Torts and Crimes," *Columbia Law Review* 85, 95-92.

Montmarquette, C. and M. Nerlove, 1985, "Deterrence and Delinquency: An Analysis of Individual Data," *Journal of Quantitative Criminology* 1, 37-58.

O'Donnell, J. A., *et al.*, 1977, *Young Men and Drugs: A Nationwide Survey*, Research Monograph No. 5, National Institute on Drug Abuse, Washington, D.C.

Parker, J. and H. G. Grasmick, 1979, "Linking Actual and Perceived Certainty of Punishment," *Criminology* 17, 366-379.

Piliavin, I., *et al.*, 1986, "Crime, Deterrence, and Rational Choice," *American Sociological Review* 51, 101-119.

President's Commission on Law Enforcement and Administration of Justice, 1967, *The Challenge of Crime in a Free Society*, U.S. Government Printing Office, Washington, D.C.

Pyle, D. J., 1983, *The Economics of Crime and Law Enforcement*, St. Martin's Press, New York.