Journal of Economic Perspectives—Volume 10, Number 1—Winter 1996—Pages 43–67

Crime, Punishment, and the Market for Offenses

Isaac Ehrlich

The persistence of illegal activity throughout human history and some of its apparent regularities have long attracted the attention of economists. For example, Adam Smith (1776 [1937], p. 670) observed that crime and the demand for protection from crime are both motivated by the accumulation of property. William Paley (1785 [1822]) presented a penetrating analysis of factors responsible for differences in the actual magnitudes of probability and severity of sanctions for different crimes. Jeremy Bentham, the father of utilitarianism, focused considerable attention on the calculus of both offenders' behavior and the optimal response by the legal authorities.

It was not until the late 1960s, however, that economists reconnected with the subject, using modern economic analysis.[1] In this paper I shall focus on two of the main themes that characterize the literature on crime in the last three decades. The first is the evolution of a "market model" that offers a comprehensive framework for studying the problem. Like the classical approach, the model builds on the assumption that offenders, as members of the human race, respond to incentives. Of course, not every single offender does so. But willful engagement in even the most reprehensible violations of legal and moral codes does not preclude an ability to make self-serving choices, and this has been the justification for applying economic analysis to all illegal activities, from speeding and tax evasion to murder.

[1] The seminal paper is Becker (1968). Other works that led the resurgence of the interest in crime include Fleisher (1966), Tullock (1967), Rottenberg (1968) and the papers by Ehrlich, Landes, Posner and Stigler collected in Landes and Becker (1974).

■ *Isaac Ehrlich is Melvin H. Baker Professor of American Enterprise and Leading Professor of Economics, State University of New York at Buffalo, Buffalo, New York.*

Indeed, the distinguishing feature of the major contributions by economists has been the attempt to explain the various aspects of crime through the tools of optimization and equilibrium analysis, rather than by reliance on deterministic social and environmental factors that are independent of the human will. At least in the economic literature, there has been little controversy concerning this approach.

The second theme concerns a more controversial issue: what constitutes an optimal crime control policy. The economic paradigm recognizes two sets of incentives that motivate potential or actual offenders: negative and positive. Negative incentives are those that deter and otherwise prevent would-be and actual offenders from entering or actively pursuing illegitimate activity: the probability and severity of punishment, and the type of punishment to be imposed. Positive incentives are those that induce participation in legitimate alternatives to crime: legitimate employment and earning opportunities, rehabilitation programs and a lower disparity in the distribution of income in society. The controversy that has emerged in the literature—subtle in some cases, explicit in others—concerns the relative efficacy and desirability of negative versus positive incentives and thus the appropriate policy inferences to be drawn. I shall try to evaluate some basic points of contention in this controversy.

As a starting point, it may be useful to look at some background data concerning the problem of crime. Consider Tables 1 and 2. The crime rate figures in Table 1, as well as the growing number of offenders in prisons shown in Table 2, imply that crime has been a growth industry in the United States over the last few decades.[2] At the same time, Table 1 also shows that both the probability and severity of punishment for specific crimes have generally been falling over the last three decades. A lower percentage of offenses known to the police is resulting in an arrest; the probability of imprisonment is smaller; and the time served in prison is shorter. The growth in the prison population, substantial as it is, has not kept up with the even larger growth in criminal behavior.

The Market for Offenses

What I call the "market model" of crime is based on five key assumptions which are typical of economic theory in general. First, offenders, potential victims, buyers of illegal goods and services, and law enforcement authorities all behave in accordance with the rules of optimizing behavior. Second, they generally form

[2] The crime rates in Table 1 are those reported by the Uniform Crime Reports of the FBI, which are based on complaints of crime by victims. A puzzling issue is that the trends of some of the FBI-reported crime categories appear to be quite different from those reported by the National Crime Victimization Studies (NCVS), which are based on statistical surveys of the incidence of victimization. The two sets of data cannot be compared over the entire period covered by Table 1, however, because the NCVS series are not available before 1973.

Table 1
Crime and Law Enforcement Indicators for Index Crimes[a], 1960 and 1991

	Murder and Non-negligent Manslaughter	*Forcible Rape*	*Robbery*	*Aggravated Assault*	*Burglary*	*Larceny*	*Auto Theft*	*All Index Crimes*
[b]Crime Rate (per 100,000)								
1960	5.1	9.6	60.1	86.1	508.6	1034.7	183.0	1887.2
1991	9.8	42.3	272.7	433.3	1251.8	3228.5	658.9	5897.3
[c]Median (Mean) Time Served in State Prisons Before First Release, in Months								
1960	52.0[b]	30.0[c]	33.9	19.5	20.4	16.7	18.9	NA
	(121.4)	(44.8)	(42.4)	(25.0)	(24.6)	(19.8)	(21.3)	(28.4)
1991	68.0	44.0	27.0	15.0	15.0	9.0	11.0	17.1
	(84.0)	(56.0)	(40.0)	(22.0)	(22.0)	(14.0)	(14.0)	(24.1)
[d]Percentage of Offenses Known to Police Cleared by Arrest								
1960	92.3	72.5	38.5	75.8	29.5	20.1	25.7	30.8
1991	67.2	51.8	24.3	56.5	13.5	20.3	13.9	21.3
[e]Probability of Imprisonment (State Prisons)[3]								
1960	39.8[e]	22.7	8.4	3.0	2.4	2.2	2.1	2.8
1991	28.4	5.2	3.6	1.7	1.0	0.2	0.4	0.8

Notes: [a] Index crimes include the seven categories listed below, as defined in the Uniform Crime Report.
[b] 1960 median time is for homicide, including negligent manslaughter.
[c] 1960 median time is for all sex offenses.
[d] Percentage of those entering state prisons relative to offenses known.
[e] 1960 probability is for homicide, including negligent manslaughter.
Sources (by row): Sourcebook of Criminal Justice Statistics, 1992, Table 3.122. Data are based on revised Uniform Crime Reports. Characteristics of State Prisoners, 1960, Table A1, and R3. National Corrections Reporting Program, 1991, Table 2.3. Uniform Crime Reports 1960 and 1991. Uniform Crime Reports 1961 and 1991, Characteristics of State Prisoners, 1960, National Corrections Reporting Program, 1991.

expectations about relative legitimate and illegitimate opportunities, including severity and certainty of punishment, based on available information, so that subjective expectations and objective opportunities can be linked. Third, there is a stable distribution of preferences for crime, as well as for safety from crime, in the population. Fourth, since crime is an external diseconomy by definition, and public law enforcement is a prime example of a nonexclusionary public good, the objective of law enforcement is generally presumed to be maximization of social welfare. Fifth, aggregation conditions concerning the behavior of all relevant parties assure well-defined equilibria. These assumptions lead to an equilibrium model of crime.

By the "market" for offenses, I do not mean necessarily a physical setting where illegitimate transactions are contracted, but the more abstract notion of a Walrasian

Table 2
Prisoners in State and Federal Institutions

	1979	*1986*	*1991*
Total Prisoners in State and Federal Institutions	301,470	522,084	789,347
Total Prisoners in State Institutions	274,564	450,416	711,643
Percentage of Index Crime Offenders	73.8	68.3	56.6
Percentage of Drug Offenders	6.4	8.6	21.3

Sources: Bureau of Justice Statistics, *Sourcebook of Criminal Justice Statistics,* 1992, Table 6.58 and 6.70. U.S. Department of Justice, Bureau of Justice Statistics Bulletin, "Prisons and Prisoners," 1982.

market in which the aggregate behavior of suppliers and demanders is coordinated and made mutually consistent through adjustments in relevant prices. In Becker (1968), equilibrium is achieved just through the interaction between offenders and law enforcers. In reality, other parties are involved as well: consumers or patrons of illicit goods and services in specific crimes, and potential victims. These parties determine the direct or derived demand for specific illegitimate activities. Public intervention "taxes" the incidence of crime through its interaction with both the supply and demand sides of the market.[3]

The Supply of Offenses

A person's decision to participate in illegal activity i can be viewed as motivated by the costs and gains from such activity. These include the expected illegitimate payoff (loot) per offense, w_i; the direct costs incurred by offenders in acquiring the loot (including the costs of self-protection to escape punishment), c_i; the wage rate in an alternative legitimate activity, w_l; the probability of apprehension and conviction, p_i; the prospective penalty if convicted, f_i; and finally one's taste (or distaste) for crime—a combination of moral values, proclivity for violence, and preference for risk. For analytical simplicity, assume that offenders pursue only a single criminal or legitimate activity.

A straightforward combination of these components into an overall expected net return per offense, π_i, might read that this is equal to expected gross payoff − direct costs incurred in acquiring the loot − foregone wages from legitimate activity − (probability of conviction) × (prospective penalty if convicted).[4] For crimes that do not involve any material gain, the net return is negative; it can be viewed as the price of crime to the offender.

[3] The following analysis extends Ehrlich (1981), in which the concept of the market is introduced and applied. For related analyses and illustrations, see Van den Haag (1975, ch. 5), Balkin and McDonald (1981) and Cook (1986).

[4] For those who prefer notation to words, this condition would be $\pi_i = w_i - c_i - w_l - p_i f_i$.

This formulation includes two additional simplifying assumptions. One is that potential offenders are risk neutral; the other is that "distaste for crime" can be measured as a constant, compensating expected net return that an individual requires to enter a criminal activity. In other words, the net payoff must exceed some threshold level before an individual decides to engage in crime. Given these conditions, the individual supply of offenses will be a function of the personal (expected) net return from crime.

This way of thinking about an individual's decision to participate in crime offers some powerful insights about the aggregate flow of offenses. For example, imagine that potential offenders faced identical legitimate and illegitimate opportunities. The shape of the market supply-of-offenses schedule would then be determined by the distribution of ethical values in the population as reflected by the different thresholds. These would dictate the *minimal* net returns that different individuals would require before entering a criminal activity. If the distribution were bell-shaped, the supply of offenses per capita, or the crime rate q_i, would be an increasing function of the *actual* net return per offense, as illustrated by the inflected and upward-sloping schedule *SS* in Figure 1.[5]

Herein lies the first lesson to be derived from the economic approach to crime: even if individual supply-of-offenses functions were completely inelastic with respect to variations in net returns above their critical threshold levels, so that active offenders would not react to either positive or negative incentives above these levels, it is still true that the market supply curve would be generally elastic. This is because changes in the actual net return from crime would make the latter exceed or fall below the threshold level of marginal offenders, thus inducing the latter to enter into or exit from criminal activity.

More realistically, one can extend this supply analysis to allow for varying legitimate market wages for different persons, but assume a stable distribution of these wages about the mean wage. In this case, the market supply curve should be interpreted as a function of the *average* net return per offense, and its shape would depend on the joint probability distribution of individuals' taste for crime and their legitimate wages.[6] Similarly, the analysis could allow for differences in individuals' perceptions of their probability of apprehension and punishment about its expected value—say, the current average probability.[7]

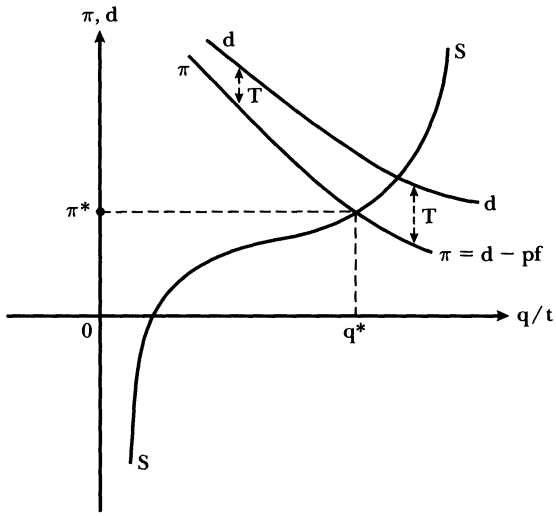
It is plausible to assume that community wealth, which is a major determinant of the gross payoff in crimes involving material gain, is exogenous to the model. Similarly, the distribution of legitimate wages may be largely exogenous, because it

[5] That is, $q_i = S(\pi_i)$, with $S'(\pi_i) \geq 0$.

[6] In the case of "crimes of passion" with no material payoff, where the net reward from crime is therefore negative $\pi_i = -c_i - w_i - p_i f_i$, the supply-of-offenses function would have to be drawn entirely within the southeast quadrant of Figure 1. But the general proposition concerning aggregate supply behavior applies here as well: an increase in the price of crime—a reduction in π_i toward a greater negative value—would lower its quantity supplied.

[7] A process of expectation formation by individuals concerning these probabilities is postulated by Sah (1991).

Figure 1
The Market for Offenses



is determined in the economy-wide labor market of which the criminal sector is a small segment. However, the direct cost of crime to the offender, the probability of conviction and the penalty if convicted are affected by private and public actions to achieve protection from crime, and thus are endogenous to the model. The next two subsections consider the latter activities.

Private Protection and the Demand Side of the Market

There is literally a direct demand for activities defined as crime in the case of illegal goods and services, including stolen goods, which are purchased by willing consumers. But even in the case of violations of persons and property, there is a derived demand for crime, or a tolerance of crime, that is inversely related to the demand for private protection. Individuals desire protection from crime, but protection comes at a cost. Potential victims can purchase insurance policies, or they can affect both the probability of becoming victims and the extent of their loss if victimized by employing a myriad of safety measures ranging from locks, burglar alarm systems and safety deposit boxes, to paying higher rents for living in safe neighborhoods or hiring private guards—a combination of self-insurance and self-protection (Ehrlich and Becker, 1972).

Because protection is costly, its optimal amount would be set at a level where its marginal costs and benefits are in balance. An expected-loss minimizing solution can be shown to imply that the optimal expenditure, e_i^*, would be directly related to the perceived risk of victimization, v_i, and the prospective private loss from crime, L_i. The intuitive reasoning is that individuals are expected to increase or lower their

"defenses" against the threat of victimization in some proportion to its expected severity. A factor of proportionality, a, which we take to be a constant, reflects the productivity of self-protective efforts in increasing the offender's direct cost of crime, c_i, as well as the latter's discouraging effect on the probability of victimization.[8] Assuming rational expectations on the part of potential victims, the average perceived risk of victimization, $v(e_i^*)$, for the representative person is the same as the crime rate in the population, q_i. We can thus conceptualize a direct demand-for-protection schedule as an increasing and concave function of the crime rate.

How do we get from the direct demand for protection to the derived-demand-for-offenses schedule—*dd* in Figure 1? The latter schedule shows why a given crime rate, or risk of victimization, is "demanded" (tolerated) by potential victims as an indirect consequence of what they are willing to spend to reduce their risk of victimization.[9] Higher expenditures on protection raise the time and effort offenders must spend on acquiring the potential loot from their victims, c_i. Any increase in the latter, in turn, means a reduction in the differential "gross" return per offense over the direct and opportunity cost incurred by the offender (the loss of a legitimate wage), in the absence of any sanctions, d_i.[10] The bottom line is that a higher crime rate, or risk of victimization, q_i, induces greater efforts at self-protection by potential victims, which in turn reduce the differential gain to offenders, d_i. Hence the downward-sloping shape of the derived-demand curve *dd*.[11]

Public Enforcement

Since crime is, by definition, an externality, and the maintenance of law and order is essentially a public good, the economic literature has focused mainly on the determination of optimal means of law enforcement and crime control, rather than the basic rationale for public rather than private enforcement of laws.[12] Three main issues have been at stake. One is the absolute and relative magnitudes of probability versus severity of punishment; for example, is it more productive to have many convictions with relatively smaller punishments or fewer convictions with relatively longer punishments? Should sanctions be applied uniformly, or adjusted by the severity of the offense and the prior record of the offender? A second issue is which penalties are more efficient: monetary fines, imprisonment, or probation, restitution and other "intermediate punishments." The third issue focuses on the

[8] An algebraic formulation of this condition would read $e_i^* = a_i v(e_i^*) L(e_i^*)$, where a_i denotes the sum of the elasticities of v and L with respect to e.

[9] That is, the demand law can be described by $d_i = D(q_i)$, with $D'(q_i) \leq 0$.

[10] The algebraic formulation would be $d(e_i) = w_{(i)} - w_{(l)} - c(e_i)$.

[11] Another reason for the downward-sloping shape of the derived demand for offenses are potential scale diseconomies at the market level. Because criminal targets are available in limited quantities in any given geographical (market) area, offenders would tend to select at any point in time the targets that yield the highest potential return per time spent. Thus, if the number of offenders increases in a given area due to factors unrelated to the available stock of criminal targets, the marginal offenses would be associated with less-remunerating targets.

[12] For thoughts about the role of privately provided law enforcement, see Landes and Posner (1975) and Friedman (1984).

relative usefulness of deterring would-be offenders as opposed to the incapacitation and/or rehabilitation of known offenders.

The approach economists have taken toward these choices has generally been based on a "public interest" criterion: the law enforcement authority seeks to maximize social welfare by minimizing the losses from crime, including the costs of law enforcement and crime control. The specification of the relevant social welfare function, however, involves normative as well as positive considerations.

We shall return to the welfare criteria issue, but for the sake of a simple closure, assume that the enforcement objective is to minimize the per capita social cost of crime through an optimal expected sanction, $T_i = p_i f_i$, which combines the probability of arrest and conviction and the penalty if convicted. This sanction must be produced at some positive cost of enforcement, $E(p_i f_i, q_i)$, subject to due process constraints.[13]

The optimal public expenditures on law enforcement and the resulting expected sanction, or "tax" on crime, T_i, are determined in roughly the same way optimal private protection was determined—that is, by balancing marginal costs of enforcement with marginal benefits of crime prevention. But there is one important difference. Private crime prevention considers only the potential private costs of crime. Public enforcement, in contrast, is a monopolized state activity, and is therefore determined in view of its marginal efficacy in reducing the equilibrium crime rate and social loss from crime. Moreover, since public enforcement is applied to all criminal activities, any interactions among markets for different crime categories (because of substitutability and complementarity among specific offenses) or across different geographical areas (because of spillover effects in crime control) must also be accounted for in arriving at an optimal enforcement strategy.[14]

The marginal reduction in the equilibrium amount of crime depends not just on the productivity of law enforcement in raising the expected sanction and the latter's effect on the supply of offenses, but also on the elasticity of the private demand schedule, or on possible shifts in the supply schedule resulting from incapacitative sanctions or rehabilitative programs (effects discussed later in this paper). As a starting point, however, it is simpler to abstract from incapacitation and rehabilitation, and focus instead on the case of sanctions that serve only to deter

[13] The function $E(pf,q)$ incorporates both the direct costs of law enforcement and the indirect costs of punishing offenders. The production function of law enforcement activity implicit in the former cost component recognizes the crowding effect which a higher volume of offenses may exert on the probability and even severity of punishment at a given enforcement budget (Becker, 1968; Ehrlich, 1973), but since budgets are assumed to be determined optimally, this crowding effect is internalized in setting optimal values of p and f.

[14] The internalization of relevant external effects for each crime and across crime categories produces a set of propositions concerning the optimal magnitudes of probability and severity of punishment—most notably that the severity of expected sanctions must "fit the crime"—that are generally consistent with their ranking in Table 1 (Becker, 1968; Stigler, 1970). For additional insights concerning optimal magnitudes of probability and severity of sanctions under more general social welfare functions see, for example, Polinsky and Shavell (1979) and Ehrlich (1982).

potential criminals. In this case, the supply and (derived-) demand schedules already discussed are independent of the public "tax schedule" on crime.

Given the relevant market parameters, the optimal expected punishment $p_i f_i$ can be shown to be a generally increasing function of the crime rate (Ehrlich and Gibbons, 1977).[15] It is shown as the implicit "tax" schedule TT in Figure 1, which accounts for the vertical distance between the private demand schedule dd and the net return (or "after-tax") schedule, $\pi\pi$.[16]

Market Equilibrium

The supply of crime, together with the demand for private and public protection from crime, as shown in Figure 1, form the basic components of the market for specific offenses. This market will be in equilibrium when the quantity of crime, q^* in Figure 1, is such that neither criminals (looking at the net expected return from crime), private individuals (looking at their risk and costs of victimization), nor government (looking at the relevant social welfare function) find it necessary to adjust their behavior and alter the prevailing net return or price associated with crime.

This "market setting" has important implications concerning the themes of this paper. The model implies (as does Durkheim, 1958) that crime is a "normal" social fact which is assured of historical survival at some positive level regardless of the prevailing economic, political or social system. Crime persists in this model because the interplay between private supply and demand forces and the social costs of enforcement imply that some level of crime will remain socially optimal—that is, tolerable. Similarly, recidivism by known offenders is also an implication of the model rather than an exception to it, because offenders are expected to discount the actual risks of apprehension and punishment in deciding to participate in crime. Private self-protection and public law enforcement set a "price," or "tax," on criminal activity by reducing the marginal net return to the offender. The existence of equilibria in the market for offenses suggests that these prices are effective, at least on the margin.

Of course, this conclusion does not mean that social, political or demographic conditions are irrelevant. Instead, these factors are captured by the components of the expected returns from illegitimate and legitimate pursuits, as determined by both supply and demand forces. Extensions of the model under dynamic conditions seek to explain the role of the family, education, life-cycle factors and social interactions in determining legitimate and illegitimate opportunities, and thus the dynamic stability of supply and demand schedules in specific markets.[17]

[15] That is, $p_i f_i = T(q_i, \psi)$, with $T'(q_i) \geq 0$, where ψ is a vector summarizing the relevant market parameters.

[16] This illustration ignores possible complementarities in production between public and private protective activities. Thus, as long as punishment involves a purely deterring sanction, both the supply and demand schedules in Figure 1 are independent of the public "tax schedule."

[17] For some current work along these lines, see Glaeser, Sacerdote and Scheinkman (1995) and Ehrlich and Lui (1995). Also see Ehrlich (1975b).

This model can account for apparent secular trends in crime without resorting to assumptions about changes in "taste" for criminality. Economic growth and real asset accumulation raise the potential payoff to the offender (w_i) in many criminal activities. If the distribution of legitimate earning opportunities and private and public protection efforts remain constant, then the rise in wealth or affluence would shift upward the derived-demand schedule *dd*, and thus the net-demand schedule $\pi\pi$, along the supply schedule *SS* in Figure 1, resulting in more crime.[18] Supreme Court rulings on the scope of offenders' rights to due process exert a largely exogenous influence on the productivity of law enforcement efforts, regardless of the basic merit of these rulings as safeguards for civil and human rights. Changes in Court decisions or sentencing guidelines over time shift the implicit tax schedule *TT*, and thus the net return schedule $\pi\pi$ in Figure 1. They partly explain the observed trends in measures of the probability of arrest, conviction and punishment, as well as in the equilibrium flow of offenses. Similar considerations on both the demand and supply side of the market for offenses can explain acute differences in actual crime rates across poor and affluent neighborhoods without resorting to the possibility of multiple equilibria in the market for offenses.[19]

Limits on the Power of Positive and Negative Incentives

Limiting Supply and Demand Responses

In the simple market model of the preceding section, equal changes in negative and positive incentives produce the same change in the expected net return from crime, and therefore the same absolute deterrent effect on the marginal offender. This conclusion does not follow, however, once we allow for nonneutral attitudes toward risk. It can be shown that a 1 percent increase in the probability of punishment will have a greater deterrent effect than a 1 percent increase in the severity of punishment on the decision to enter an illegal activity if the marginal offender is a risk preferrer, while the opposite is the case for a risk avoider (Becker, 1968).

If offenders engage in crime on a part-time basis, however, as is the case empirically (U.S. Department of Justice, 1991; Reuter et al., 1990), there is a

[18] Higher wealth may also increase private and public protection against crime, as observed by Adam Smith, but even if all the components of the net return per offense rise by the same proportion as w_i, its absolute value, π_i, would also rise by the same proportion, leading to more crimes involving material gains. Even murder may increase as a result (despite the expected increase in private demand for self-protection and life saving), since this crime is often committed as a byproduct of robbery and other felonies. Indeed, the proportion of all murders where victims and offenders are unrelated persons (as opposed to those involving family members, friends and acquaintances) has increased drastically from 23 percent in 1960 to 53 percent in 1991, according to the Uniform Crime Report.

[19] Analyses leading to multiple equilibria—for example, Freeman, Grogger and Sonstelie (1995)—typically assume, however, that both private and public expenditures on crime control are exogenously determined constants.

theoretical possibility that negative and positive incentives would be ineffective at the individual level because of conflicting income and substitution effects. For example, if offenders prefer risk, then the loss of potential income caused by a higher average penalty per offense, by itself, may induce them to shift the allocation of their working time toward more crime: at a lower income position a risk preferrer is more willing to gamble the cost of punishment for the chance to obtain a higher criminal return (Ehrlich, 1974). This kind of ambiguity applies more generally to the effects of negative incentives if time spent in illegal pursuits were an inferior good (Block and Heineke, 1975). But however important these ambiguities may be at the individual level, they are likely to disappear at the market level because changes in illegitimate incentives do not generate income effects for marginal entrants.

More important, the market framework implies that it is not enough to look only at aggregate supply or demand reactions to changes in incentives. In analyzing the equilibrium quantity and price in the market for a particular good, it makes a considerable difference whether supply is shifting for a good that is inelastically or elastically demanded. Similarly, in the market for offenses, the equilibrium level of crime will depend on interactions between supply and demand forces. Take, for example, the effect of exogenous changes in "general" incentives, such as increases in probability and severity of fines or increases in the average legitimate wage. Such changes would lower the net return from crime, and thus the net demand curve $\pi\pi$ in Figure 1. But the effect on the equilibrium volume of crime would now depend on the shape of both the supply of offenses and the derived demand for crime schedules, which are dictated by reactions of offenders, potential victims, and law enforcers to changes in incentives.

General incentives, such as monetary fines and legitimate market wages, operate on the population at large, including active offenders. The importance of distinguishing individual from market-level responses becomes even greater in relation to the effects of "specific incentives," such as incapacitative penalties and rehabilitation programs, which are targeted on convicted offenders. The efficacy of such specific incentives in preventing crime is constrained by the fraction of offenders who are apprehended and punished and their actual period of confinement in the case of incapacitative penalties, and by the even smaller fraction of released offenders who will remain rehabilitated over their labor careers, in the case of rehabilitative programs.

The successful removal of incarcerated or rehabilitated offenders from the illegitimate market causes a leftward shift in the supply-of-offenses schedule, but the effect on the equilibrium volume of offenses is now *inversely* related to the elasticity of the aggregate supply curve. The reason is that "removed" offenders can be replaced by either new entrants or an intensified illegal activity on the part of active offenders. Suppose, for example, that the market supply of illicit drugs were infinitely elastic. The mere removal of some active drug dealers from the marketplace, or an occasional drug bust, would then cause no change in transactions, if the demand schedule remained intact.

It is plausible that certain types of offenders are more easily replaced than others. For example, sellers of illegal goods and services are probably replaced more easily than those who commit violent crime. The effectiveness of incapacitation or rehabilitation for sellers of illegal goods and services, therefore, would be highly limited, but such programs could be more efficacious for perpetrators of violent crimes.

Incapacitative penalties, such as imprisonment, cause both a deterrent effect on those at large and an incapacitative effect on those imprisoned. Rehabilitation programs, in contrast, even if entirely successful in eliminating recidivism by released offenders, do not exert a deterrent effect. Indeed, the training or work subsidy implicit in successful job-oriented programs has a potential counter-deterrent effect as it reduces the prospective penalty costs for would-be offenders. Incapacitation and imprisonment without any attention to rehabilitation, however, could increase the odds of recidivism by released offenders because of the informal training for illegitimate activities they receive from associating with other convicts.

The general theme that is emerging here is that while specific incentives may be effective at the level of individual offenders, the market model points to significant limitations of their potential efficacy in reducing the aggregate flow of many crimes. By the same token, even if general incentives have a weak effect at the individual level, they may be efficacious at the market level because of the impact they have on the entry and exit of marginal offenders.[20]

Limits Imposed by Optimal Enforcement Policies

If the objective of public law enforcement is to maximize social income—that is, to minimize the aggregate (or per capita) social loss from crime—it follows that law enforcement must be effective on the margin. The reason is that probability and severity of punishment entail positive marginal costs, which would be optimal to incur only if probability and severity of punishment effected a marginal reduction in crime. If higher sanctions increase only the indirect social cost of punishment, but not the direct cost of arresting and convicting offenders, the aggregate loss-minimizing criterion also requires that the absolute elasticity of offenses with respect to the probability of punishment exceed that with respect to its severity (Becker, 1968). The reason is that the costlier it is on the margin to increase the means of enforcement, the larger must be its marginal benefit, as indicated by its marginal deterrent and preventive effect.[21] But this restriction does not necessarily hold if the aim of law enforcement is to maximize a social welfare function that incorporates a concern for distributive outcomes of law enforcement in addition to income maximization, as discussed further in the penultimate section of this paper.

[20] For evidence on the efficacy of rehabilitation at the individual offenders' level, see Cook (1975), Lipton, Martinson and Wilks (1975), Rossi, Berk and Lenihan (1980) and Piehl (1994). Estimates of the incapacitative effect of imprisonment are provided in Ehrlich (1981) and Levitt (1995b).

[21] The elasticities must also be less than unity in this case because otherwise law enforcement will lower not just the frequency of offenses, but even the number of convicted offenders and will thus not be costly on the margin.

Perhaps the sharpest restriction on the efficacy of law enforcement instruments, which is an outcome of optimal choices by both offenders and law enforcers, is the prediction that the absolute elasticity of crime with respect to the probability of apprehension will exceed that with respect to the conditional probability of conviction given apprehension, and that the latter would exceed the elasticity with respect to the conditional probability of a more severe punishment given conviction (Ehrlich, 1975a). The rationale on the supply side is that the more general the event leading to undesirable consequences for the offender, the greater is the deterrent effect associated with its probability. Optimal enforcement justifies this same result because raising the probability of apprehension, for example, requires not just raising the costs of apprehending offenders, but also the costs of convicting and punishing them in later stages of the enforcement process. The deterrent effects of these probabilities must then have the same ranking as their relative costs.

Estimates of the Effects of Positive and Negative Incentives

The empirical literature concerning the effects of positive and negative incentives on crime is voluminous; for example, see the surveys by Palmer (1977), Pyle (1983), Freeman (1983) and Cameron (1988). Taken as a whole, these studies offer a mountain of evidence consistent with the hypothesis that both negative and positive incentives have a deterrent effect on crime. The evidence comes from studies using alternative crime reports; different crime categories and population groups; different countries, states, cities and even city tracts (Thaler, 1977); time series, cross-section and panel data; and both aggregate and individual samples. Moreover, a number of studies find evidence consistent not just with the model's implications concerning the direction of effects of deterrence variables, but also with the latter's relative magnitudes, as predicted by the elasticity conditions above. While most of the evidence is derived from regression estimates of supply-of-offenses functions, corroborating results are obtained from comparisons of "market" returns associated with alternative illegitimate activities, which differ in the degree of punishment risk they impose on offenders (Viscusi, 1986).

This evidence is important to followers of the market model because the model may be used to study alternative crime control policies only if incentives do indeed matter. By contrast, if pure deterrence were estimated to be weak relative to, say, the force of incapacitation, then locking up convicted offenders would appear to be the only option available for effective law enforcement.

Yet it would be premature to view the supportive empirical evidence as conclusive. My main concern is not that a minority of studies fail to find a deterrent effect of either negative or positive incentives: this is predictable on probabilistic grounds alone, even when the same econometric models are being implemented. But while the qualitative results obtained from most studies are similar, the quantitative estimates vary. Such differences may be attributable, in part, to the use of data from different levels of aggregation: for example, the effect of employment opportunities

on crime across cities may be quite different from their effects across states, or for the nation as a whole over time, because of different spillover effects across markets, or because cross-sectional (as opposed to cyclical) variations in employment opportunities reflect shifts in supply as well as demand conditions in the labor market. Also, as noted earlier, even if individual data produced estimates of inelastic supply-of-offenses schedules, the aggregate schedule could still be quite elastic.

More troublesome is that differences in econometric estimates may be the result of incomplete, or even inconsistent, specification of the market model, which are exacerbated by the intrinsic limitations of crime statistics. The following sections explore some of the common issues that arise in attempting to implement the market model, as well as their bearing on the second theme of the paper: the role of negative and positive incentives.

Specification

Since the empirical evidence developed in the literature on crime and deterrence comes mainly from regression analyses of statistical data, or "uncontrolled experiments," a meaningful estimation of the hypothesized incentive effects requires careful econometric specification of the model being tested. The task involves not just the specification of the supply-of-offenses function, but at least an implicit specification of the complete market model, because of the simultaneous relationship between crime and law enforcement. As the market model indicates, supply responses to changes in alternative incentives can be identified econometrically only if changes in these incentives come about as a result of shifts in demand schedules, such as dd or $\pi\pi$ in Figure 1, while variables controlling the location of the supply schedule are held constant.

The problem has been recognized from the outset. The standard specification in many studies of aggregate data (for example, Ehrlich, 1973; Carr-Hill and Stern, 1973; Phillips and Votey, 1975) has been a three-part econometric structure: 1) supply-of-offenses functions relating specific crime rates to measures, or proxies, of each of the law enforcement products and legitimate and illegitimate wage variables discussed earlier; 2) production functions relating the same law enforcement products to resources expanded in their production and factors affecting resource productivity, including the crime rate itself; and 3) expenditure functions, relating enforcement outlays to the crime rate, measures of per capita (social) losses from crime and underlying political constraints that may cause public expenditures to adjust to their desired level with some lag.

Typically missing from this specification, however, are indicators of derived-demand functions—that is, measures of self-protection by potential victims—which affect the differential payoffs per offense.[22] The absence of private-protection variables implies that what has been estimated empirically are relationships between

[22] Interesting exceptions are Vandaele's (1978) estimation of a partial market system for auto theft, Bartel's (1975) and Clotfelter's (1977) studies of the demand for private protection, and Goldberg and Nold's (1980) analysis of crime reporting as self-protection by victims.

equilibrium crime rates and indicators of incentives that reflect the elasticities of both supply and demand functions. These estimated responses are the relevant ones from a policy perspective, since law enforcement does not control private protection efforts. Also, their relative magnitudes need not be affected because all are subject to the same demand-side adjustment. However, the absolute levels of the estimated responses may understate the pure supply elasticities associated with all incentives.

Also, a complete specification of supply functions for specific crime categories requires the introduction of probability and severity of sanctions for both the estimated categories and interrelated crimes, or a joint estimation of interrelated crime categories using seemingly unrelated estimation procedures. For example, violent crimes may have the attributes of complementary activities—some murders are committed in the course of a robbery—whereas burglary and larceny may be substitutes. Some studies report both complementarity and substitutability across specific offense categories (Ehrlich, 1977; Levitt, 1995b).

Ideally, the functional specification of relevant structural equations should be derived from the underlying objective functions of the participants in the market model. This task has so far proved elusive, especially in connection with supply-of-offenses functions. But typical assumptions concerning the structure of reporting errors in aggregate data and the hypothesized effects of measures of costs and rewards to crime point toward the use of a logarithmic specification, which also permits direct estimation of supply elasticities. These considerations generally receive support from standard tests for optimal functional specifications (Ehrlich, 1977; Layson, 1983). Studies using less efficient specifications in analyzing aggregate data are therefore less likely to confirm the existence of deterrent effects.

Errors of Measurement

The FBI's commonly used Uniform Crime Reports (UCR) data on crime are subject to notorious underreporting errors, largely because reporting a crime is costly to the victim. Moreover, some measures of the probability of punishment are computed as ratios of the number of convictions or prison commitments to the volume of reported offenses. Errors of measurement in the latter may impart a negative bias to regression estimates of the deterrent effect of the probability measure and also bias the estimated effects of other explanatory variable that are correlated with it (Ehrlich, 1974, pp. 127–30).

There are, however, standard approaches for dealing with these problems. As an approximation, reporting errors are likely to be proportional to the true crime rate, so the logarithms of the reported crime figures and probability measures can be thought of as proxies for the true variables. Another standard remedy for potential biases from random reporting errors involves using instrumental variables to estimate a "predicted" probability measure that is in principle free of such errors. A more direct approach involves modeling the economics of reporting on the part of victims and integrating it in the econometric model to distinguish reported from "true" crime rates.

Yet another approach is to test the model against alternative sources of crime data, such as National Crime Victimization Surveys (NCVS) and vital statistics data (in connection with homicide). UCR and NCVS data give quite different pictures of the time trend of specific offenses, especially violent crimes, although recently the differences have narrowed. Deterrent effects of private and/or public protection measures are confirmed in these studies as well (Goldberg and Nold, 1980; Ehrlich, 1977).

Measurement errors in reported crime may not operate in one direction only: victims may be more likely to report a crime when the probability or severity of punishment is higher or if they receive compensation. In regressions using reported crime rates, therefore, measures of criminal sanctions may also be subject to a positive spurious correlation with crime rates. In a study relevant to this point, Levitt (1994) finds the net effect on reporting errors of both police presence and the standard measure of probability of arrest to be rather small and statistically insignificant.

Selecting Empirical Counterparts of Theoretical Constructs

It can be difficult to measure the theoretical variables of the market model with readily available data. For example, punishment is never quantified in terms of its cost to the offender. Even if time served in prison by currently released offenders is a good predictor of time to be served by current convicts, the discounted present value of the future cost of imprisonment is not proportionally related to its length.

Similar problems affect the measurement of positive incentives for avoiding crime, like the level of legitimate earnings and employment opportunities. For example, using time series variations in the rates of unemployment and labor force participation to measure such opportunities captures mainly cyclical fluctuations in labor-demand conditions, unlike their cross-sectional counterparts, which are influenced by labor supply decisions as well. It is not surprising that time series measures prove to have a weaker link with crime than do cross-sectional measures at a point in time.

Some studies may use average legitimate earnings in the population as a measure of positive incentives, but not only may this average correlate poorly with the legitimate earnings of active offenders and marginal entrants, it is usually strongly correlated with median community income, or wealth, which is a proxy for illegitimate earning opportunities. One remedy has been to introduce both median income and the percentage of families below one-half of the median income to account for the disparity between illegitimate and legitimate earnings of actual and potential offenders. Both variables have been found to produce sizable and statistically significant effects on the incidence of property crimes, and even on the incidence of murder, partly because murder can be a byproduct of robbery (Ehrlich, 1973, 1977).

An alternative remedy involves the use of individual data where information is available on arrests and convictions per months free, legitimate earnings, the fractions of past arrests resulting in conviction, and the severity of past sentences. Indeed, studies based on such data, with some exceptions, provide a more direct support for the deterrent effect of both legitimate earnings and the conditional

probability and severity of punishment (Witte, 1980; Myers, 1983; Trumbull, 1989; Grogger, 1991; Tauchen, Dryden and Griesinger, 1994).

One problem with studies employing individual data to estimate supply functions is the use of a person's incidence of arrest or conviction during the sample period as a measure of that person's crime rate. But arrests are a product of the individual-specific crime rate and concurrent probability of arrest. The past probability of being convicted or the severity of the imposed punishment can be expected to affect (in different degrees) both the person's crime rate and probability of arrest in subsequent periods, as the offender's identity is exposed to both potential victims and law enforcement agents. The estimated deterrent effects of certainty and severity of punishment (and for similar reasons that of legitimate earnings) do not provide, therefore, direct estimates of the elasticities of individual crime rates with respect to the chance of being arrested or the length of sentence. Nor can this approach produce conclusive inferences about the relative efficacy of certainty vs. severity of punishment.

Identification Restrictions

The most serious econometric challenge has been to assure that the estimated relationship between crime rates and measures of law enforcement reflect the causal effect of the latter on the former, rather than vice versa. In terms of the market model, the econometric structure must contain some exogenous or predetermined variables that are included in the demand function for law enforcement or in the production functions of law enforcement measures, but are excluded from the supply-of-offenses function. Changes in these exogenous variables will shift the implicit "tax" and net-demand schedules in Figure 1 and allow us to identify the shape of the supply-of-offenses schedule.

Data concerning such exogenous or predetermined variables that are excluded from the supply function are hard to come by. Some proxies used in a number of studies on specific crimes include past levels of total expenditure on police and courts, past crime rates and measures of urbanization and population density, which constrain the productivity of resources spent on law enforcement. In many studies, severity of punishment is measured by using the length of time served by those released from state prisons; since this figure tended to be constant over previous decades, it has been treated as a predetermined variable. The results of studies using such identification restrictions tend to confirm the existence of discouraging effects on crime of both positive and negative incentives (Ehrlich, 1973, 1975a; Carr-Hill and Stern, 1973; Phillips and Votey, 1975; Wolpin, 1978a).

These identification restrictions are subject to potential weaknesses.[23] In particular, crime rates may be serially correlated, and some demographic variables used

[23] See the critiques in Blumstein, Cohen and Nagin (1978). These critiques focus exclusively on potential negative biases in the estimates of deterrent effects because of the crowding effects exerted by high crime rates on both probability and severity of imprisonment (see note 10), although biases are likely in the opposite direction as well because of optimal enforcement.

to identify the supply-of-offenses function may also affect the supply functions themselves. But serial correlation can be accounted for in time series regressions, and identification restrictions are testable. With some exceptions, studies pursuing the relevant statistical tests and using the estimation techniques justified by these tests have actually found strong statistical support for the impact of positive and negative incentives (Layson, 1983, 1985; Ehrlich and Brower, 1987).

Some studies introduce per capita expenditures on law enforcement activity as a direct determinant of the structural supply of offenses, in lieu of measures of probability and even severity of punishment, even though the earlier analysis of public decision making indicates that such expenditures are likely to be simultaneously determined by, and positively associated with, the crime rate. Without pursuing the proper identification restrictions, it is no wonder that some of these studies find a weak association between, say, expenditures on police and crime.

A promising approach toward the identification problem is to quantify political or institutional (judicial) factors that affect either law enforcement budgets or the "rules of the game" that influence the productivity of enforcement efforts in determining the probabilities of apprehension and conviction. Since law enforcement budgets are determined by elected officials at the state and local levels, they may be subject to electoral cycles. Election dates and Supreme Court rulings are unlikely to be affected by concurrent crime rates. Their impact on enforcement spending and productivity can therefore be used to identify the relevant supply responses. Recent studies that use such institutional and political factors have found new support for the hypothesis that law enforcement instruments do discourage crime (Ehrlich and Brower, 1987; Levitt, 1995a).[24] Studies of individual data that combine individual and area statistics on law enforcement expenditures reach similar conclusions (Tauchen, Dryden and Griesinger, 1994).

Separating Deterrence from Incapacitation

Perhaps the most unique econometric challenge has been to separate the pure deterrent effect from the incapacitation effect inherent in imprisonment. Separation of the two effects is critical for establishing the validity of the market model.

A decomposition of the total effect of imprisonment on crime into deterrent and incapacitative components can be achieved by using sample data to provide estimates of the maximum possible effect of incapacitation, based on theoretical considerations, and then comparing these to regression estimates of the actual effect of imprisonment, based on the same sample (Ehrlich, 1981). An alternative decomposition method has been used by Levitt (1995b), based on identifying a pure substitution effect of the probability of arrest for one crime on the incidence of a "substitute" crime. Both studies conclude that deterrence constitutes the dominant effect of criminal sanctions.

[24] Alternative identification procedures, based on time series techniques, have been pursued by Phillips and Ray (1982) in their analysis of murder. They too find support for the deterrence hypothesis, including the deterrent effect of capital punishment.

Some studies have attempted to assess the deterrent effects of criminal sanctions directly, by studying the impact of monetary fines which in principle, can exert only a pure deterrent effect. Use of monetary fines in felony crimes has been highly restricted in the United States, but some evidence consistent with a deterrent effect of fines on specific property crimes has been derived using United Kingdom data (Wolpin, 1978a). Evidence on the deterrent effect of monetary fines has also been inferred from U.S. data concerning antitrust violations (Block, Nold and Sidak, 1981).

Inferences About Positive and Negative Incentives

Freeman (1983) observes that most econometric studies using both aggregate or individual data find negative deterrent effects of both probability and severity of criminal sanctions, but that the estimated deterrent effects of positive incentives, such as employment and legitimate earnings measures, are weaker by comparison. This should not be interpreted to suggest that positive incentives are unimportant.

Since positive and negative incentives are generally correlated, consistent estimates of the separate effects of either set of incentives cannot be obtained without accounting for the influence of both. Yet many studies focusing on positive incentives—using such variables as unemployment, labor force participation, income or earnings measures and demographic variables—have completely ignored measures of the probability and severity of punishment. Freeman (1983) reports that 70 percent of the studies of positive incentives he had surveyed are subject to this omission. Similarly, some studies focusing on the role of criminal sanctions omit measures of positive incentives. Others include only estimates of probability of arrests but not of the (conditional) probability of punishment or its actual severity.[25] As noted earlier, some studies that have attempted a more complete implementation of supply-of-offenses functions find statistically significant deterrent effects associated with proxies of both positive and negative incentives.

The present evidence does not allow one to conclude that positive incentives are either less or more potent than negative ones. The effects of positive incentives that concern the general population can be expected to differ from those applying to marginal offenders, and the effects of rehabilitation and employment-incentive programs that are targeted at imprisoned or ex-offenders are expected to have substantially different impacts at the individual, relative to the aggregate-market level. Some of the ambiguous effects of unemployment and legitimate earnings indicators can be ascribed to the incentive measures used rather than to the pure effect of incentives. Also, the elasticities of crime rates with respect to (general)

[25] Such omissions are common especially in the early sociological literature on deterrence (for example, Sellin, 1959, 1967). Leamer (1983) and McMannus (1985) propose to sort out the merits of incorporating either set of incentives in the supply-of-offenses function through "extreme bounds analysis," which assigns no weight to systematic theoretical considerations. They claim that results are sensitive to researchers' prior beliefs, but their tests have been shown to rely on vacuous test statistics (McAleer, Pagan and Volker, 1985; Ehrlich and Liu, 1995).

positive and negative incentives are not comparable, even if the two had the same absolute (pure) deterrent effects on individual would-be offenders, as the simplified model underlying Figure 1 suggests.

The empirical evidence seems more consistent in connection with the deterrent effects of the probability of apprehension relative to the conditional probability of conviction or the conditional probability of a larger punishment, as implied by the sharp elasticity conditions discussed earlier (Ehrlich, 1975a, 1977; Wolpin, 1978a). As expected, efforts at apprehending and convicting offenders are found to be a stronger deterrent to crime than efforts at achieving a specific sentence. Also, many studies find that increasing the risk of imprisonment for most crime categories has a significantly larger deterrent effect in elasticity terms than increasing the length of imprisonment, especially for violent crimes, and that the magnitudes of the elasticities is less than 1. The evidence is consistent with Becker's (1968) proposition about the ranking of these elasticities, which was based on a social cost minimizing enforcement strategy. I cannot yet conclude, however, that severity of punishment has a substantially weaker effect than its certainty.

On a theoretical level, others have developed models where the latter elasticity condition need not hold (see note 14). These broader models imply that the marginal cost of raising the severity of punishment might exceed that of raising the probability of punishment. Higher penalties may also raise the direct cost of prosecuting and convicting offenders because they induce greater self-protection by offenders. In both cases, optimal law enforcement implies that in equilibrium, the marginal preventive effect of severity of punishment in elasticity terms might even exceed that of its certainty.

There is also a problem with measuring severity of punishment empirically. In some of the studies reporting smaller elasticities of crime with respect to severity of punishment, severity is measured in terms of the length of sentence rendered, rather than the actual time served (for example, Grogger, 1991). As noted earlier, even measuring severity by the actual time served fails to discount its future cost, and thus treats, say, a 10-year term as twice as costly as a five-year term. This bias worsens for crimes with longer imprisonment terms.

Perhaps the single most debated issue in the literature on crime has been the deterrent effect of capital punishment. As argued at the outset, the economic approach to crime is expected to apply to all illegitimate activities, including murder and other "crimes of passion," regardless of whether these entail any material benefits. Violent offenders can still be expected to respond to an increase in the threat of punishment imposed on violent behavior, especially when the potential sanction is capital punishment. Influential studies by sociologists have concluded that capital punishment was not a deterrent to murder (Sellin, 1959, 1967), but these studies did not account for the effects of the complete set of "prices" expected to affect behavior on the margin—the probability of apprehension and conviction, the conditional probability that the death penalty be actually imposed, the severity of the alternative punishment of imprisonment and other relevant incentives. Since murder is also the best-reported crime, the results of testing the

deterrence hypothesis in connection with this crime can provide inferences about the validity of the market model in general. There is now corroborating evidence from several studies using independent time series and cross-section data from the United States, Canada and the United Kingdom that is consistent with the hypothesis that punishment in general, and capital punishment specifically, have a deterrent effect on murder (Ehrlich 1975a, 1977; Layson, 1983, 1985; Wolpin, 1978b; Phillips and Ray, 1982).

The econometric issues raised in the debate about the validity of these findings are the same as those already discussed above.[26] What may be worth stressing is that, as predicted, in all studies supporting the deterrence hypothesis, the conditional risk of execution, while having a significant deterrent effect, has the least impact on the incidence of murder relative to equal percentage changes in apprehension, conviction or punishment risks, and that the alternative, and most frequently used sanction for murder—the length of imprisonment—also exerts a statistically significant discouraging effect on murder. Lower levels of wealth, income inequality and unemployment (in time series studies) have also been found to deter the incidence of murder.

Some Policy Implications

The relative desirability of specific means of crime control cannot be determined just by their relative efficacy; it also depends on their relative social costs and on the welfare criteria invoked as a justification for public law enforcement.

For example, if the welfare objective is to maximize social income, then the social cost of purely deterring sanctions, such as fines, would be close to zero, because as transfer payments, fines are free of the deadweight costs associated with imprisonment, house arrests, probation and other intermediate punishments. An optimal enforcement strategy may then involve raising such fines to their maximal feasible level (consistent with a convict's wealth constraint) while lowering the probability of apprehension and conviction to its minimal level (Becker, 1968). Even under this (narrow) efficiency criterion, however, it would be optimal to use imprisonment and intermediate punishments along with fines for those crime categories where the added incapacitation value of imprisonment justifies its added costs.

The enforcement strategy would be different if the social welfare function were broadened to include distributional objectives as well. These include, for example, a preference for promoting equality of individuals under the law, reducing the legal error of convicting the innocent, or lowering the corollary prospect of letting the

[26] For example, see the critiques and studies by Blumstein, Cohen and Nagin (1978), Avio (1979), Hoenack and Weiler (1980), and the responses by Ehrlich and Mark (1977), Ehrlich and Brower (1987) and Layson (1985).

guilty go free. For example, since the probability of apprehension and punishment is substantially less than one, penalties are in fact applied through a lottery system. Offenders who are caught and punished are subjected to ex post discrimination under the law because they "pay" not just for their own crime, but also for offenders who get away with crime. The degree of such discrimination rises as the penalty becomes more severe, or if the probability of punishment is very low. Such concerns help explain why severity of punishment is often traded on the margin for a higher probability of apprehension and conviction. It also helps explain why the justice system introduces numerous safeguards to protect the rights of the accused, and why the opposition to capital punishment tends to increase when the penalty is applied infrequently and capriciously (Ehrlich, 1982).

Incorporating concerns for equality and legal error in the social welfare function raises not just the marginal social cost of severity of punishment, but that of any strategy of enforcement (as long as the probability of being arrested and punished for a crime is low) relative to its cost under the narrower efficiency criterion. The implication is that more crime would be tolerated as a result of a tradeoff between equity and efficiency in enforcement—a tradeoff typical of social choice in general.

This analysis is applicable to crime control strategies concerning the use of positive incentives as well. The market model implies that a lower disparity in the distribution of earning opportunities in legitimate markets will deter offenders on the margin, by reducing their differential gains from criminal activity. This provides a justification for public policies aimed at equalizing educational and employment opportunities partly as means of reducing crime. However, since these policies, unlike conventional law enforcement, cannot be targeted specifically at actual or potential offenders, they may entail relatively high social costs as means of crime control.

The positive implications of the market model and some corroborating empirical evidence concerning the relative efficacy of deterrence vs. incapacitation and rehabilitation, as discussed earlier in this paper, suggest a direction of reform of the criminal justice system through greater reliance on general incentives and purely deterring sanctions. Forcing offenders to pay fines through work-release programs (including direct restitution to their victims) may in many cases be as effective a means of crime prevention as the more costly incapacitating penalties or rehabilitative programs—especially in the case of many theft crimes or transactions in illicit goods and services. Thus, the dramatic growth in the proportion of those imprisoned for drug offenses, shown earlier in Table 2, appears to be inconsistent with this implication of optimal enforcement.

Conclusion

The market model of crime is still a work in progress. Data limitations have so far precluded a complete implementation of its relevant structure, and the model

itself is evolving. But the existing literature on the whole lends strong support to the basic premises of the model.

A common confusion about the deterrence hypothesis is that it applies only to negative incentives, while positive incentives may hold a greater promise for "solving" the crime problem. Another often-heard claim is that we don't need to know more about punishment because punishment does not eliminate crime. Both claims are wrong. The deterrence hypothesis and its logical extension—the market model—rely on the *marginal* efficacy of *both* positive and negative incentives and on the interaction between market demand and supply forces, to explain the observed variability in the frequency of offenses across space and time. The empirical evidence is consistent with the hypothesis that punishment and other general incentives exert a deterrent effect on offenders. This suggests, for example, that there is no need to rely exclusively on harsh or incapacitating sanctions to achieve efficient crime control. A better understanding of what does works, however, calls for more, rather than less, research into the general deterrence hypothesis and the market model based on it.

■ *I wish to thank Chaya Ehrlich and the editors of the journal, especially Timothy Taylor, for valuable editorial assistance and suggestions. I have also benefitted from comments by Gary Becker, Richard Posner, Govind Hariharan and Zhiqiang Liu. Shao-Chi Chang and Jianguo Zhong have provided valuable assistance.*

References

Avio, Kenneth, "Capital Punishment in Canada: A Time-Series Analysis of the Deterrence Hypothesis," *Canadian Journal of Economics,* November 1979, *12*:4, 647–76.

Balkin, Steven, and John F. McDonald, "The Market of Street Crime: An Economic Analysis of Victim-Offender Interaction," *Journal of Urban Economics,* November 1981, *10*:3, 390–405.

Bartel, Ann, "An Analysis of Firm Demand for Protection Against Crime," *Journal of Legal Studies,* June 1975, *4*:2, 443–78.

Becker, Gary, "Crime and Punishment: An Economic Approach," *Journal of Political Economy,* 1968, *76*:2, 169–217.

Becker, Gary, and William Landes, *Essays in the Economics of Crime and Punishment.* New York: Columbia University Press, 1974.

Block, Michael, and John Heineke, "A Labor Theoretical Analysis of Criminal Choice," *American Economic Review,* June 1975, *65*:3, 314–25.

Block, Michael, Frederick Nold, and Joseph Sidak, "The Deterrent Effect of Antitrust Enforcement," *Journal of Political Economy,* 1981, *89*:3, 429–45.

Blumstein, Alfred, Jacqueline Cohen, and Daniel Nagin, *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates.* Washington, D.C.: National Academy of Science, 1978.

Cameron, Samuel, "The Economics of Crime and Deterrence: A Survey of Theory and Evidence," *Kyklos,* 1988, *41*:2, 301–23.

Carr-Hill, R. A., and Nicholas Stern, "An Econometric Model of Supply and Control of Recorded Offenses in England and Wales," *Journal of Public Economics,* November 1973, 2, 289–318.

Clotfelter, Charles, "Public Service, Private Substitutes and the Demand for Protection Against Crime," *American Economic Review,* December 1977, *67*:5, 867–77.

Cook, Philip, "The Correctional Carrot: Better Jobs for Parolees," *Policy Analysis,* Winter 1975, *1,* 12–51.

Cook, Philip, "The Supply and Demand of Criminal Opportunities." In Tonry, Michael, and Morris Norval, eds., *Crime and Justice: An Annual Review of Research.* Vol. 7, Chicago: University of Chicago Press, 1986, pp. 1–28.

Durkheim, Emile, *The Rules of Sociological Methods.* 8th ed., New York: Free Press, 1958.

Ehrlich, Isaac, "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy,* 1973, *81*:3, 521–65.

Ehrlich, Isaac, "Participation in Illegitimate Activities: An Economic Analysis." In Becker, Gary, and William Landes, eds., *The Economics of Crime And Punishment.* New York: Columbia University Press, 1974, pp. 68–134.

Ehrlich, Isaac, "The Deterrent Effect of Capital Punishment: A Question of Life and Death," *American Economic Review,* June 1975a, *65*:3, 397–417.

Ehrlich, Isaac, "On the Relation Between Education and Crime." In Juster, F. T., ed., *Education, Income and Human Behavior.* New York: McGraw-Hill, 1975b, pp. 313–37.

Ehrlich, Isaac, "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence," *Journal Political Economy,* 1977, *85*:4, 741–88.

Ehrlich, Isaac, "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation and Deterrence," *American Economic Review,* June 1981, *71*:3, 307–22.

Ehrlich, Isaac, "The Optimum Enforcement of Laws and the Concept of Justice: A Positive Analysis," *International Review of Law & Economics,* June 1982, *2,* 3–27.

Ehrlich, Isaac, and Gary Becker, "Market Insurance, Self-Insurance and Self-Protection," *Journal of Political Economy,* 1972, *80*:4, 623–48.

Ehrlich, Isaac, and George Brower, "On the Issue of Causality in the Economic Model of Crime and Law Enforcement: Some Theoretical Considerations and Experimental Evidence," *American Economic Review,* May 1987, *77*:2, 99–110.

Ehrlich, Isaac, and Joel Gibbons, "On the Measurement of Deterrent Effect of Capital Punishment and the Theory of Deterrence," *Journal of Legal Studies,* January 1977, *6*:1, 35–50.

Ehrlich Isaac, and Zhiqiang Liu, "Extreme Bounds Analysis and the Deterrence Hypothesis—Where is the Con?," mimeo, 1995.

Ehrlich, Isaac, and Francis Lui, "Corruption and Economic Growth," mimeo, State University of New York at Buffalo, 1995.

Ehrlich, Isaac, and Randall Mark, "Fear of Deterrence: A Critical Evaluation of the 'Report of the Panel on Research on Deterrent and Incapacitation Effects,'" *Journal of Legal Studies,* June 1977, *6*:2, 293–316.

Fleisher, Belton, *The Economics of Delinquency.* Chicago: Quadrangle, 1966.

Freeman, Richard, "Crime and Unemployment." In Wilson, J. Q., *Crime and Public Policy.* San Francisco: Institute for Contemporary Studies, 1983, pp. 89–106.

Freeman, Scott, Jeffrey Grogger, and Jon Sonstelie, "The Spatial Concentration of Crime," mimeo, 1995.

Friedman, David, "Efficient Institutions for the Private Enforcement of Law," *Journal of Legal Studies,* June 1984, *13*:2, 379–97.

Glaeser, Edward, Bruce Sacerdote, and Jose Scheinkman, "Crime and Social Interaction," mimeo, 1995.

Goldberg, Itzhak, and Frederick Nold, "Does Reporting Deter Burglary?—An Empirical Analysis of Risk and Return in Crime," *Review of Economics and Statistics,* August 1980, *62,* 424–31.

Grogger, Jeffrey, "Certainty Versus Severity of Punishment," *Economic Inquiry,* April 1991, *29,* 297–309.

Hoenack, Stephen, and William Weiler, "A Structural Model of Murder Behavior and the Criminal Justice System," *American Economic Review,* June 1980, *70*:3, 327–41.

Landes, William, and Richard Posner, "The Private Enforcement of Law," *Journal of Legal Studies,* January 1975, *4*:1, 1–46.

Layson, Stephen, "Homicide and Deterrence: Another View of the Canadian Time-Series Evidence," *Canadian Journal of Economics,* February 1983, *16*:1, 52–73.

Layson, Stephen, "Homicide and Deterrence: A Reexamination of the United States Time-Series Evidence," *Southern Economic Journal,* July 1985, *52,* 68–89.

Leamer, Edward, "Let's Take the Con Out of Econometrics," *American Economic Review,* March 1983, *73,* 31–43.

Levitt, Steven, "Reporting Behavior of Crime Victims and the Size of the Police Force: Implications for Studies of Police Effectiveness Using Reported Crime Data," mimeo, Harvard University, 1994.

Levitt, Steven, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." NBER Working Paper Series No. 4991, 1995a.

Levitt, Steven, "How Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?," mimeo, Harvard University, 1995b.

Lipton, D., R. Martinson, and J. Wilks, *The Effectiveness of Correctional Treatments: A Survey of Treatment Evaluation Studies.* New York: Praeger, 1975.

McAleer, Michael, Adrian Pagan, and Paul A. Volker, "What Will Take the Con Out of Econ-

ometrics?," *American Economic Review,* June 1985, *75,* 293–307.

McManus, Walter, "Estimates of the Deterrent Effect of Capital Punishment: The Importance of the Researcher's Prior Beliefs," *Journal of Political Economy,* April 1985, *93,* 417–25.

Myers, Samuel, "Estimating the Economic Model of Crime: Employment Versus Punishment Effects," *Quarterly Journal of Economics,* February 1983, *98,* 157–66.

Paley, William, *The Principles of Moral and Political Philosophy.* 1785. Reprint, London: T. Davidson, Whitefriars, 1822.

Palmer, John, "Economic Analysis of the Deterrent Effect of Punishment: A Review," *Journal of Research in Crime and Delinquency,* January 1977, *14*:1, 4–21.

Phillips, Llad, and Subhash Ray, "Evidence on the Identification and Causality Dispute about the Death Penalty." In *Applied Time Series Analysis.* Amsterdam: North-Holland, 1982, pp. 313–40.

Phillips, Llad, and Harold Votey, "Crime Control in California," *Journal of Legal Studies,* June 1975, *4*:2, 327–49.

Piehl, Anne, "Learning While Doing Time," mimeo, Working Paper Series R94–25, Kennedy School, Harvard University, 1994.

Polinsky, Mitchell, and Steven Shavell, "The Optimal Tradeoff Between the Probability and Magnitude of Fines," *American Economic Review,* December 1979, *69*:5, 880–91.

Pyle, David, *The Economics of Crime and Law Enforcement.* London: Macmillan, 1983.

Reuter, Peter, Robert MacCoun, and Patrick Murphy, "Money from Crime: A Study of the Economics of Drug Dealing in Washington, D.C.," reviewed by Henry Miller, *Contemporary Drug Problems,* Winter 1990, *18*:4, 713–6.

Rossi, Peter, Richard Berk, and Kenneth Lenihan, *Money, Work, and Crime: Experimental Evidence.* New York: Academic Press, 1980.

Rottenberg, Simon, "The Clandestine Distribution of Heroin, Its Discovery and Suppression," *Journal of Political Economy,* 1968, *76*:1, 78–90.

Sah, Raaj, "Social Osmosis and Patterns of Crime," *Journal of Political Economy,* 1991, *99*:6, 1272–95.

Sellin, Thorsten, *The Death Penalty.* Philadelphia: American Law Institute, 1959.

Sellin, Thorsten, *Capital Punishment.* New York: Harper & Row, 1967.

Smith Adam, *The Wealth of Nations.* 1776. Reprint, New York: Random House, 1937.

Stigler, George, "The Optimum Enforcement of Laws," *Journal of Political Economy,* 1970, *78*:3, 526–36.

Tauchen, Helen, Ann Dryden, and Harriet Griesinger, "Criminal Deterrence: Revisiting the Issue with a Birth Cohort," *Review of Economics and Statistics,* August 1994, *76*:3, 399–412.

Thaler, Richard, "An Econometric Analysis of Property Crime: An Interaction Between Police and Criminals," *Journal of Public Economics,* August 1977, *8,* 37–51.

Trumbull, William, "Estimation of the Economic Model of Crime Using Aggregate and Individual Level Data," *Southern Economic Journal,* October 1989, *56,* 423–39.

Tullock, Gordon, "The Welfare Costs of Tariffs, Monopolies, and Theft," *Economic Inquiry,* June 1967, *5,* 224–32.

U.S. Department of Justice, Bureau of Justice Statistics, *National Corrections Reporting Program.* Washington, D.C.: U.S. Department of Justice, 1991.

U.S. Department of Justice, Bureau of Justice Statistics, *Sourcebook of Criminal Justice Statistics.* Washington, D.C.: U.S. Department of Justice, 1992.

U.S. Department of Justice, Federal Bureau of Investigation, *Uniform Crime Reports.* Washington, D.C.: U.S. Department of Justice, 1960, 1961, and 1991.

U.S. Department of Justice, Federal Bureau of Prisons, *Characteristics of State Prisoners.* Washington, D.C.: U.S. Department of Justice, 1960.

Vandaele, Walter, "An Econometric Model of Auto Theft in the United States." In Heineke, J. M., ed., *Economic Models of Criminal Behavior.* New York: North-Holland, 1978, pp. 303–90.

Van den Haag, Ernest, *Punishing Criminals.* New York: Free Press, 1975.

Viscusi, Kip, "The Risks and Rewards of Criminal Activity: A Comprehensive Test of Criminal Deterrence," *Journal of Labor Economics,* 1986, *4*:3, 317–40.

Witte, Ann Dryden, "Estimating the Economic Model of Crime with Individual Data," *Quarterly Journal of Economics,* February 1980, *94,* 57–84.

Wolpin, Kenneth, "An Economic Analysis of Crime and Punishment in England and Wales," *Journal of Political Economy,* 1978a, *86*:5, 815–40.

Wolpin, Kenneth, "Capital Punishment and Homicide in England: A Summary of Results," *American Economic Review,* May 1978b, *68*:2, 422–27.