URBAN PROPERTY CRIME AND THE

DISTRIBUTION OF INCOME

By

LONNIE K. STEVANS

Bachelor of Science
Hofstra University
Hempstead, New York
1978

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1980

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1983

DEDICATION

My wife, Vicky,

Sons, Zachary and Michael

URBAN PROPERTY CRIME AND THE

DISTRIBUTION OF INCOME

Thesis Approved:

Keith Olson

Thesis Adviser

Orley M. Amos Jr.

Lyle Broemeling

Gerald M. Lage

Norman N. Durham

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to express my gratitude to Frank and Dori Boccio and Paul Tricarico for their help in data collection which made the empirical work of this study possible.

TABLE OF CONTENTS

Chapter | Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

To evaluate personal welfare, members of society refer to their relative as well as their absolute incomes. When these relative comparisons are inconsistent with a group's sense of equity, conflicts in society are likely to emerge. Economic growth with a constant distribution of income leads to rising individual incomes, as well as a rising social standard. Thus, more income for the individual will mean increased happiness only if everyone's income has not similarly risen (Easterlin, 1973). The purpose of this dissertation is to focus on the process by which relative welfare comparisons produce one type of conflict--crime. Specifically, urban property crime will be viewed as an aggregate consumption externality associated with the distribution of income.

Objective

Criminal activity may be represented by the following multiplicative form:

$$C = C_1(X)\ C_2(\rho)\ C_3(\vec{E}).$$

X - a measure of the income distribution

ρ - the probability of punishment

$\vec{E}$ - a vector of random determinants

It is assumed that potential criminals respond to all factors so that "C" is never zero. The first term, "$C_1(X)$". is the amount of crime that results because of relative income comparisons. The second term is based on the assumption that deterrence affects crime, through the probability of being arrested and convicted. The third, "$C_3(\vec{E})$", accounts for the other determinants of crime, both psychological and sociological.

The main objective of this study is to focus on that component of crime which is related to the distribution of income "$C_1(X)$". If people view the distribution of income from an individual perspective, then any perceived change in this distribution becomes a necessary condition for the committal or non-committal of a criminal act. The income distribution may therefore exhibit a certain degree of "publicness" in consumption. To the extent that this is true, the standard pareto income-leisure conditions will result in more criminal activity than is socially optimally, since no one takes into account the effect that individual incomes have on the income distribution and crime.

This study differs from previous economic analyses of crime in essentially two ways. First, a theoretical model is presented which specifies relative income comparisons as an important determinant of criminal behavior. Second, criminal activity is viewed as an aggregate consumption externality or an "incidental" effect resulting from an income distribution which is not Pareto optimal. Crime affects everyone adversely, not just the victims, through a fear of being victimized. Many studies on the economics of crime have alluded to this external effect, although the major emphasis has been on the independent choice aspect ("$C_2(\rho)$" above).

From a public policy perspective, the existence of interdependent preferences implies an additional option for reducing crime: redistribution. The role of the local government is to determine how to optimally reduce the level of crime--by choosing that option which has the highest benefit-cost ratio. In this case, I assume that there are only two options: income redistribution or deterrence. The proposed empirical formulation in this analysis examines the viability of these two alternatives as a measure of reducing criminal activity by a comparison of the relative magnitudes of their cost-effectiveness. The test does not assume the existence of any conscious public policy to reduce crime through transfers. Rather, it is simply a means to determine which option, redistribution or deterrence, has been most cost-effective. The study is undertaken for the New York City area.

Outline

Chapter II of this study will synthesize some recent work on the economics and sociology of crime, and pareto optimal redistributions. Chapter III will provide a theoretical framework which hypothesizes that the distribution of income plays a significant role in explaining the level of criminal activity. An empirical model is developed in Chapter IV, and results of the analysis are presented in Chapter V. Finally, Chapter VI will review the findings and provide some suggestions for future research.

CHAPTER II

BEHAVIORAL PERSPECTIVES OF CRIMINAL ACTIVITY

Crime is a major social concern which has received increasing attention in recent years. This chapter will analyze criminal behavior from a socio-economic perspective, since an adequate understanding of crime requires a multidisciplinary approach.

To the economist, crime is rational behavior--a choice that is made by a person or persons in deciding how best to spend their time. In making the choice, individuals consider what they stand to gain and what they stand to lose; that is, they consider the benefits and costs of using their time in different ways--working legally, working illegally, or not working at all. An additional implication is that individuals have some knowledge, not necessarily perfect, of the benefits and costs associated with different actions. Because of the assumption of rationality, economists would hypothesize that economic decisions are the <u>only</u> necessary conditions for the committal of a criminal act. Therefore, any policy that would increase costs relative to benefits would deter crime. This does not imply that in every instance in which punishment increases one can expect to see crime decrease. There are other factors which also affect crime rates, such as material costs, psychic costs, etc. But economists would argue that, "ceteris paribus," an increase in expected punishment costs should reduce crime (Tittle, 1973). Finally, of course, there may be instances in which the assumption of rationality is not appropriate, for example, "crimes of passion."

To some sociologists, crime is considered to be deviant behavior. The deterrent effect of punishment concerns the role of sanctions in generating conformity. Much of the empirical research on the question of the ability of punishment to deter crime has been done by sociologists and socio-psychologists, although from a different perspective than economists, and perhaps sometimes with different expectations regarding the effectiveness of punishment. Sociologists may also focus on other aspects of criminal behavior, such as the motivation behind the deviant behavior and the impact of class, race, or sex (Horton, 1973).

Some of the concerns of sociologists and other social scientists who study criminal behavior have been incorporated within the economic model. The gains and costs of criminal behavior include psychic elements which is a "catch-all" for all kinds of psychological, sociological, and political phenomena. Most economic models of crime do not emphasize or examine these phenomena, but they are nevertheless included and typically viewed as "exogenous" determinants--much like tastes and preferences in consumer theory. On the subject of criminal motivation and class conflict, for example, it has been found that crime is related to the distribution of income, but such interactions have never been explained adequately in the economic literature on crime (Ehrlich, 1975). Consider the following example: one possible gain to a low-income ghetto resident from committing armed robbery may be the criminal's "psychic" benefits from achieving a more favorable position on the income scale. Since this act is included in the average gain from crime, all effects as a result of the committal of a criminal act are normally thought to be reflected in the market for crime. A moment of thought, however, should reveal this approach to be incorrect. First, the existence of a relative income comparison, or a concern for distribution, implies interdependent preferences. Thus, utility functions will be altered by a changing distribution and all of the relevant gains and losses will not be accounted for by the market. Second, crime affects everyone, not just the parties involved. One does not have to be a victim in order to be adversely affected by crime.

The point of departure for the approach taken in this study from the orthodox view is the inclusion of the income distribution as a significant determinant of crime. Moreover, simce all relevant third party effects are not reflected in market adjustments, criminal behavior will be viewed as an aggregate externality associated with the distribution of income; "aggregate" because the identity of the individuals committing offenses are not important, only the overall level of crime.

The Economics of Crime

It is possible to distinguish between two different perspectives adopted in economic analyses of criminality. One approach stresses the role of preferences and free will in the analysis of choice and views the criminal as a self-interested, rational decision-maker. Differences in attitudes toward risk play a central role in explaining variations in crime, while environmental or institutional factors are virtually ignored by this approach. The other approach emphasizes environment and opportunities and suggests that crimes are directly or indirectly determined by such economic factors as poverty, inequality, or the oppression of laws. Closely related to the latter approach is the set of hypotheses advanced by segmented labor market theorists who argue that though actors appear to make rational self-interested choices, their opportunities and preferences are actually determined by institutional arrangements.

The choice theoretic approaches to the analysis of crime and punishment by Becker (1968) and Ehrlich (1973) closely resemble the utilitarian writings of Beccaria (1963) and Bentham (1843). Beccaria, Simplistic thinkers for example, argued the utilitarian principle that legislation should be formed and enforced so as to assure that the greatest happiness is shared by the greatest number. Laws are useful and necessary for the security and good of society, but punishments must be introduced to deter individuals from violating society's laws.

Becker's work is based on almost identical premises as those of Beccaria: that optimal enforcement of laws should follow from the utilitarian principle of minimization of social loss, and that illegal behavior can be understood in terms of rational decision making of free-willed individuals. The scarcity of resources demands that allocative decisions within the criminal justice system follow some social welfare objective. The objective proposed by Becker, minimization of social losses or costs, is deemed reasonable because traditional policy on punishment (i.e., defining appropriate levels of the certainty and severity of social sanctions) imposes costs on and produces benefits for offenders, victims, and society.

Ehrlich (1973) expands on Becker's theory by investigating the potential criminal's optimal allocation of time to crime and work. Making choices in the face of uncertainty, the individual chooses to enter or not to enter criminal activity in the process of maximizing expected utility, calculated for contingent states of the world. Since expected utility declines for increasing certainty or severity of punishment, optimal participation in crime declines with increasing punishment.

The economic approach which places heavy emphasis on the structure of opportunities, the institutional environment in which decisions are made, and the stratification within the economy shares somewhat the deterministic perspective of many 19th century socialist criminologists. Coljanni (1954), for example, wrote that not only is crime directly related to poverty and economic misery, but fluctuations in criminal conditions are caused by disturbances in the economic situation. The Dutch criminologist, Bonger (1969), added one more causal factor to the crime equation: cupidity. Bonger argued that poverty and cupidity are fundamental to the order of capitalist society and that measures of criminality vary directly with measures of cupidity and poverty.

While few modern economists have adopted the Bonger model without qualification, a number have begun to develop an approach toward the theory of criminality which is paralleled by the development of models of segmented or stratified labor markets. The first systematic description of criminal behavior within the context of segmented labor markets was provided by Piore (1978). In reflecting on the characteristics of jobs or workers in two distinct employment sectors, Piore argues that the behavioral patterns fostered by low-paying, menial, and unpleasant "secondary labor market" jobs are reinforced by a lower-class life-style which is more compatible with welfare and illicit activity than with legitimate employment. However, these same behavior patterns--for example, lateness and absenteeism--tend to shape both the opportunities of disadvantaged workers and the characteristics of the jobs they face. In a sense, then, secondary labor market workers' actions are both determined and determining.

In summary, two different perspectives exist among economic writers on criminality: rational free-willed and deterministic. The latter emphasizes opportunities and institutional barriers while the former highlights the deterrent effect of punishment. It must be noted, however, that even though the deterministic approach emphasizes the importance of economic factors, there is no mention or analysis of relative welfare comparisons.

Critique of the Orthodox Approach

The central findings of Becker and Ehrlich have not gone unchallenged. Block and Heineke (1975) argued that the Becker-Ehrlich results are based upon restrictive assumptions about the probability distributions for success or failure in criminal activity. In general, they found that the effects of the certainty and severity of punishment on optimal participation in crime are not determinate for arbitrary success or failure distributions.

In spite of the apparent challenges on other grounds, it is interesting to note why the theorists writing in the Becker tradition have not addressed themselves to the issue of employment opportunities. First, employment policies cannot be shown to deter crime within the context of the neo-classical model. Work and crime are assumed to be substitutes. As the expected return to one activity rises, the supply of the other falls. Recognizing the possibility that either a rise in the probability of punishment or a rise in the potential wage rate may reduce participation in illegitimate activity, it does not follow, at least in the orthodox approach, that improved employment opportunities will reduce crime. One might note here that when the neo-classical model is extended to include the possibility of a backward bending supply curve for illegitimate activity, it becomes theoretically possible that increased punishment may not reduce crime either.

Secondly, the choices made to enter criminal activity are assumed to be of the same type as other economic choices, such as labor force participation, consumption, or investment in training and education. Hence, no special modeling effort is necessary in the neo-classical model to incorporate labor market behavior into a model of crime (unless there are external effects or special institutional barriers which inhibit the functioning of the market).

Evidence of Crime and Employment

Fleischer (1966) was one of the first modern economists to attempt to test econometrically the relation between employment opportunities and crime. Observed variations in crime rates were presumed to follow from changes in the demand and supply of crime. The demand for engaging in delinquent acts depended upon tastes for delinquency and on legitimate alternatives to crime behavior. The supply of delinquency was assumed to result from opportunities to commit delinquent acts. Such opportunities vary with the victim's self-protection and economic and social characteristics of the environment. Despite conceptual difficulties with his model, Fleischer was able to explain a large percent of the variation in delinquency, using both time series and cross-sectional data. Fleischer found that unemployment and mean income of the highest income quartile of families is negatively related to delinquency. This evidence is in conflict with the results of a 1958 study by Glaser and Rice (1959) in which they found that increases in unemployment were associated with increases in crime among adults but with decreases in crime among juveniles. The U. S. National Commission on Law Observance and Enforcement (1971) has found that larcenies, homicides, and imprisonment rates rise with unemployment, but violent property crimes fall. On the other hand, Friendlander (1972), using a 16-city cross-sectional sample in 1966, found crime to be negatively related to unemployment, thereby raising further uncertainties about the precise effect of employment opportunities on crime.

Economists have consistently found that economic variables are significant. Recent works have attempted to use variables such as population density, migration, income dispersion, and labor-force participation rates in explaining variations in crime rates, but with little rationalization for how there variables interact (Forst, 1973). Variables of the "economic-factors-influencing-crime" variety are usually considered in studies testing a Becker-type model. For example, measures of income inequality, wealth, and race (percent non-white) are generally positive and significant in Ehrlich's equations (see footnote 1, p. 14). However, unemployment is generally found to be insignificant.

Cook (1975), in an analysis of a sample of Massachusetts parolees, concluded that improved job opportunities reduce the probability that an ex-offender will recidivate. Taggart (1972) and others have cited findings that suggest that participation in illegal activity is linked to failure in the job market. Released offenders, however, have higher turnover rates, higher unemployment rates, and lower wages than the general population, as Pownall (1971) shows. Does this evidence suggest that training offenders or providing job counseling and referral services will reduce crime? In a review of a number of early manpower training efforts to improve the post-release labor market experiences of incarcerated offenders, Cook found that evaluations of the effectiveness of such efforts were generally inconclusive. Does this suggest that the problem is one of discrimination against ex-offenders? Cook concluded, no. He thought that the evidence from Pownall's study of ex-offenders in Baltimore and Philadelphia indicated, instead, that ex-offenders merely face the same poor employment prospects as other disadvantaged workers. The evidence should be viewed cautiously, especially in the light of Leonard's (1976) evidence suggesting that prior or current criminal record is generally regarded as grounds for dismissal from a job. Portnay (1970) has also noted similar restrictions on hiring former criminals in many skilled trades.

A resurgence of interest by the Labor Department in the post-release labor market experiences of ex-convicts has stimulated a number of studies which should generate needed data for discovering how poor employment opportunities affect participation in crime and how criminal records affect employment opportunities. Yet, early evidence from an experiment in Baltimore, wherein parolees were randomly selected to either (1) receive cash subsidies until they found employment, or (2) receive job referral services and some employment counseling and training, or (3) both, suggested that only cash subsidies significantly altered the number of repeated offenses (Lenihan, 1976).

In view of the findings to date, it seems premature to conclude that an unambiguous relationship exists between crime and punishment, and employment opportunities and crime. This is not to be considered a repudiation of prior findings of a statistical association between crime and penal measures, unemployment rates, labor force participation rates, and other economic variables. It is merely a statement of caution when viewing these findings within an analytical perspective that does not take into account the institutional setting--for instance, the urban ghetto. Moreover, much of the focus on crime from an economic standpoint has ignored the general equilibrium effects of engaging in illegal activity. It is obvious that the effects of crime on minority or ghetto communities transcend specific acts. In particular, a more general analysis of crime would reveal the interaction of participation in illegal activity and consumption patterns of both legal and illegal goods. Such an analysis, coupled with an identification of ghetto business decisions in response to crime, residential location and investment decisions, tax rates, and erosion of tax bases due to crime, would provide a basis for assessing the relative costs and benefits in ghetto communities of different anti-crime measures.

The glaring omission of discussions of general equilibrium interactions of crime, especially in urban areas, signifies a "void" in the economic literature on crime. In addition, there are some theoretical and empirical anomalies associated with both economic approaches. As mentioned previously, relative welfare comparisons cannot simply be treated as insignificant or ignored completely. The neo-classical model of crime includes these effects under the disguise of economic gains, while the segmented market approach ignores such comparisons completely--absolute measures of employment opportunities and poverty are the most important. Economic variables, such as the income distribution, have been found to have significant effects on crime, but no reason, a priori, has been given to explain the relation.[1] Many of the studies which have sought to explain criminal activity by unemployment have found conflicting results.

Such inconsistencies are due to the theoretical inadequacies of both models--neither satisfactorily handles relative comparisons. The approach to be developed in this study will analyze crime from a general equilibrium and urban perspective. Illegal activity will not be treated exclusively as the outcome of a "rational" decision-making process. It will also occur as an incidental effect resulting from the income-leisure choice of the urban affluent. Much has been said and written about growing income inequality creating feelings of malevolence, but no concise theoretical analysis has been established, at least from an economic perspective. The economic literature on Pareto optimal redistributions and the sociological literature on delinquency provide a starting point for this analysis.

---

[1]Ehrlich (1973) does confirm a positive relationship between crime rates and the degree of income inequality. However, in his analysis the median level of family income and the percentage of families with incomes below one-half the median are proxies for the returns from illegal and legal activities, respectively. These measures reflect general economic conditions, not utility interdependence, since individual utility, as specified by Ehrlich, is a function only of one's own consumption.

Pareto Optimal Redistributions

The first question to be answered is this: Given any initial distribution of income, is some redistribution necessary to achieve a Pareto optimum? Thurow (1971) gives several reasons why arbitrary initial distributions of income may not be Pareto optimal. Individuals care about the well-being of others, and may find it necessary to redistribute their income to others. Alternatively, the income distribution itself may appear in utility functions, if there are externalities involved. Achieving an optimal level of crime, for instance, may require an income redistribution that is Pareto efficient. Hochman and Rodgers (1969) argue that efficiency criteria can and should be applied to the redistribution of income through the fiscal process. The problem lies in the inability of the competitive market to generate the required outcome. Individuals may derive satisfaction from the income of others or an associated group, or preferences may exist for a particular distribution. Either way, there are incentives to avoid payment. Voluntary transfers are unlikely to achieve an optimal distribution in the first case, and where the aggregate income distribution is concerned, its properties are not unlike those of a pure public good. Exclusion is impossible, consumption is non-rival, and each individual has a vested interest in not revealing his/her preferences to avoid paying the required share of the necessary taxes.

In sum, then, it seems reasonable to assume that individuals have preferences concerning the proper distribution of income, independently from society's social welfare rankings collectively determined through the political process. It is clear, for instance, that people's behavior is sometimes motivated by altruism or envy. To the extent that this is true, the income distribution becomes a consumer externality and Pareto optimal redistributions are necessary for welfare maximization.

Development of Deviant Behavior

The final task is to examine what factors determine the pattern of tastes and preferences for crime. Much of the sociological literature on the causes of urban crime refers to concepts such as "social-class system", "inter-class conflict", "objective deprivation", etc. when analyzing illegal behavior. Sociologists, ever since Ferri (1896), have been calling attention to factors such as environment and social status. Bonger (1969) placed the blame for disproportionate crime and delinquency among the underclass on the pressures of the capitalist system. According to the Marxian point-of-view, crime is inevitable in a capitalist society. The lower classes will always commit property crimes in order to gain a more favorable position on the income scale--to succeed is imperative (Gordon, 1971). Marx himself wrote of the "working" class's desires to gain position in Bourgeois society (Marx, 1978).

Similar to the Marxian approach, but presented in a more rigorous manner is the anomie theory of Merton (1968). According to this theory, deviant behavior, at least in part, involves selective adherance to accepted social norms and occurs in areas of specific structural strains in a social system. Merton suggests that anomie develops because of a breakdown in the relationship between goals that place great stress on success, and to which all groups in our society are indoctrinated, without equivalent emphasis on institutional or legitimate channels of access to these goals. In the areas where the discrepancy between goals and means is greatest, a condition of anomie prevails, and individuals resort to illegitimate means to achieve the goals. This implies that criminal activity is, to a certain extent, "generated" by the social system.

A modification of Merton's theory appears in Cohen's (1976) theory of delinquent subcultures. This concept is also rooted in the discrepancy between goals and means. However, according to Cohen's formulation, the delinquent subculture is a reaction to socially and economically induced stresses that our social-class system inflicts on individuals. Crime results because the desire to achieve economic and social status is fulfilled by any means possible.

In an alternative, but related approach, Toby (1967) attempts to explain why crime rates are rising rapidly in affluent societies. People steal, not because they are starving, but because they are envious of the possessions of others. The rise in living standards is associated not only with an improvement of the lifestyle of the elite groups; it is associated also with the "trickling down" of television sets, automobiles, radios, etc. to segments of the population who had not anticipated such good fortune. According to Toby when expectations of more equality in the distribution of consumer goods rise faster than the standard of living, individuals will attempt to fulfill their expectations by committing property crimes.

Many theories have been advanced to account for the development of criminal behavior, and explanations for delinquent behavior have varied within and across disciplines. The sociological approach, however, expands the analysis of crime well beyond the narrow individual-centered theories that prevail in economics. Attempts by sociologists to relate delinquency to the social and economic structure have added greatly to the understanding of criminal acts as integral elements of social life, rather than exclusively as a matter of individual choice.

Conclusion

Poverty itself does not cause crime, but resentment of poverty does, and, curiously enough, resentment of poverty is more likely to develop among the relatively deprived of an affluent society than among the objectively deprived in a poor society. This is partly because affluent industrial societies are also secular societies; the distribution of goods and services here and now is a more important preoccupation than concern with eternal salvation. It is also because the mass media, to which television has been a recent but important addition, stimulate the desire for a luxurious lifestyle among all segments of the population. These considerations help to explain why the sting of socioeconomic deprivation can be greater for the poor in rich societies than for the poor in poor societies. In addition, they would also shed light on the high crime rates in many urban areas where the difference between rich and poor is even more pronounced; and on the increase in crime rates with the increase in general prosperity. Note that the positive relationship between crime and prosperity cannot be explained adequately by economic theory, alone.

In an attempt to capture the effects of deprivation, the analysis to follow will include the distribution of income as an argument in a crime function. This is not meant to be an all-encompassing approach to criminal behavior. Rather, an alternative economic model of urban property crime will be presented which maintains the concept of choice, but at the same time allows for preferences to be determined endogenously by the socially and economically induced stresses that our system inflicts on individuals.

CHAPTER III

A GENERAL EQUILIBRIUM MODEL OF CRIME

The following is an economic analysis of criminal behavior. The approach taken treats crime as an aggregate consumer externality associated with the distribution of income.

Consider a "polarized" urban area with a total population of "n" individuals, "r" rich and "p" poor:

$i = 1, 2, \ldots, p,$

$h = 1, 2, \ldots, r,$

$p+r = n.$

The poor are isolated from the rest of the urban population and unable to migrate from the ghetto to other peripheral areas.

The type of criminal activity considered in this study is victim-related or more specifically, illegal activity such as property crime. Homicides are not included, since the population of the urban area is assumed to be unchanged at any given time. Moreover, it is assumed that only a subset of the urban poor commit the crimes in question and only amongst themselves--"white collar" crimes are not considered (Cohen, 1981). Thus, in the case of property-related crimes, this would amount to a redistribution of wealth amongst the urban poor. Rising property crime rates, however, affect <u>all</u> urban residents adversely through the <u>fear</u> of being victimized. Expenditures on handguns, security devices, and public deterrence serve the purpose of minimizing this apprehension.

The objective of the government in a competitive urban economy where all choose between income and other activities, is to maximize social welfare in the context of a "first best" environment. This simply means that the planner has a set of policy tools to insure the attainment of the "bliss point" on the population's utility-possibilities frontier. All markets are competitive and behavior can be adjusted in non-competitive markets to generate the Pareto optimal results. The fundamental difference between this analysis and the other theoretical approaches to "optimal" distributions, is that the former employs an individual as well as a social view of the distribution of income: people judge their relative position on the income scale independently of society's social rankings (social welfare function). Criminal activity therefore becomes an "incidental" effect on our typical "poor" person resulting from the income-leisure choices of the "rich".

Theory

The income distribution is represented by the function "X( ):"

$$X = X(f(y_1, y_2, \ldots, y_p), g(y_1, y_2, \ldots, y_r)),$$

where X - a summary measure of the distribution of income,

f - a function representing the subdistribution of the poor,

y_i - the income of a poor individual, i = 1, 2, ..., p,

g - a function representing the subdistribution of the rich,

y_h - the income of a rich individual, h = 1, 2, ..., r.

The distribution function is assumed to have the following properties:

$$X = \begin{cases} X(f,g) & f < g \\ 1 & f = g, \end{cases} \tag{1}$$

$f > 0,\ g > 0.$

"f" and "g" are always positive, and $f \leq g$. The first and second order partials of "X" with respect to "f" and "g" are:

$$\frac{\partial X}{\partial f} > 0, \quad \frac{\partial X}{\partial g} < 0, \tag{2}$$

$$\frac{\partial^2 X}{\partial f^2} < 0, \quad \frac{\partial^2 X}{\partial g^2} > 0.$$

The shape of the function is depicted in Figures 1 and 2.

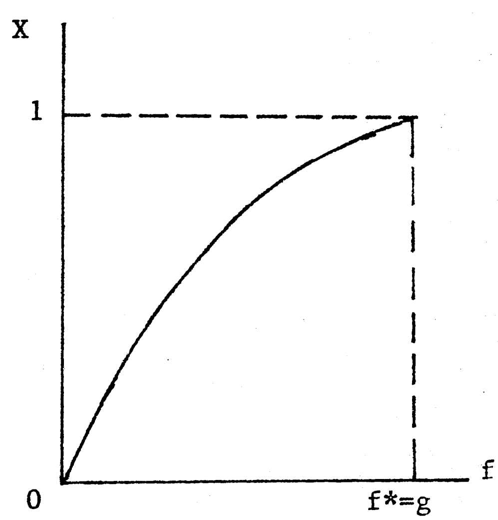

Figure 1. Distribution of f

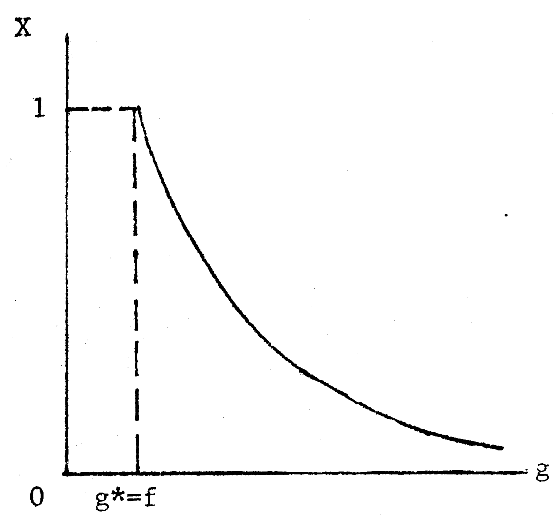

Figure 2. Distribution of g

Property (2) implies that as the poor become better-off relative to the rich ("g" constant), the distribution of income increases at a decreasing rate. Since $f \leq g$, the maximum value of "X" is one which is denoted by $f^*=g$ in Figure 1. "f" can never reach zero by the assumption $f > 0$. From Figure 2 and Property (2), it may be seen that as the rich gain relative to the poor ("f" constant), the distribution function approaches zero. Again, since $f \leq g$, the value "g*" represents the point where $X = 1$.

The third property of the function "X" is:

$$\begin{aligned} \frac{\partial X}{\partial y_i} &= \frac{\partial X}{\partial y_j}, \quad i, j = 1, 2, \ldots, p;\ i \neq j \\ \frac{\partial X}{\partial y_h} &= \frac{\partial X}{\partial y_w}. \quad h, w = 1, 2, \ldots, r;\ h \neq w \end{aligned} \tag{3}$$

Property (3) may be derived as follows. For simplicity, assume $i,j = 1,2$:

$$\frac{\partial X}{\partial y_1} = \frac{\partial X}{\partial f}\frac{\partial f}{\partial y_1}, \qquad \frac{\partial X}{\partial y_2} = \frac{\partial X}{\partial f}\frac{\partial f}{\partial y_2}.$$

A person's income matters only in that a <u>ceteris paribus</u> chance in individual income will equal the change in each subdistribution, affecting the aggregate distribution "X". Therefore:

$$\frac{\partial f}{\partial y_1} = \frac{\partial f}{\partial y_2} = 1,$$

and

$$\frac{\partial X}{\partial y_1} = \frac{\partial X}{\partial y_2} = \frac{\partial X}{\partial f}.$$

The crime function (or number of offenses committed) may now be specified as a function of the income distribution and the probability of punishment (deterrent effect):

$$C_i = C_i(X(f,g),\rho) \qquad i = 1, 2, \ldots, p \qquad C_i \geq 0 \tag{4}$$

X - Distribution function

ρ - Probability of being arrested and convicted (ρ - "rho")

In accordance with previous assumptions, crime is committed only by the poor in this model--white collar crime is not being analyzed. The probability of punishment, "ρ", depends upon two things: the probability of being arrested and charged, "ρ_a", and the probability of being convicted of the crime if arrested and charged, "$\rho_{c/a}$". To compute "ρ", one may resort to the laws of probability:

$$\rho_{c/a} = \frac{\rho}{\rho_a} ,$$

$$\rho = \rho_{c/a} \cdot \rho_a \ . \tag{5}$$

The probability of being arrested and convicted is equal to the probability of being convicted, given that the person has been arrested, times the probability of being arrested. "ρ" may be thought of as varying positively with the efficiency of the judicial system and police productivity.

The crime function has the following properties:

$$\frac{\partial C_i}{\partial y_i} = \frac{\partial C_i}{\partial X}\frac{\partial X}{\partial f} \leq 0, \qquad \left(\frac{\partial f}{\partial y_i} = 1\right) \qquad \left(\frac{\partial C_i}{\partial X} \leq 0\right) \left(\frac{\partial X}{\partial f} > 0\right) \tag{6}$$

$$i = 1, 2, \ldots, p$$

$$\frac{\partial C_i}{\partial y_h} = \frac{\partial C_i}{\partial X}\frac{\partial X}{\partial g} \geq 0 \ . \qquad \left(\frac{\partial f}{\partial y_h} = 1\right) \qquad \left(\frac{\partial C_i}{\partial X} \leq 0\right) \qquad \left(\frac{\partial X}{\partial g} < 0\right) \tag{7}$$
$$i = 1, 2, \ldots, p \ ; \ h = 1, 2, \ldots, r$$

Property (7) infers that as the urban rich become relatively better-off in the eyes of the poor, the amount of crime the poor commit amongst themselves increases. Considering the "poverty amidst plenty" which is apparent in many urban areas, as well as the influence of media advertising, this assumption of awareness on the part of the poor is not unrealistic.

If we aggregate over all poor individuals, the total amount of urban crime is obtained (some "C_i" may be zero):

$$C = \sum_{i=1}^{p} C_i = C(X,\rho) \tag{8}$$

Since each individual faces the same distribution function and probability of punishment, aggregation is similar to that of determining market demand. Given an income distribution, "X_0" and probability of punishment, "ρ_0", the total amount of crime committed would be "C_0". Increasing expenditures on crime prevention would increase "ρ" and reduce criminal activity to "C_1" (Figure 3).

To offer a more complete theory of urban crime, equation (8) and its graphical representation must be analyzed more fully. First, consider the determinants of "ρ", the probability of punishment. Expected-punishment costs may be increased by increasing "ρ". How this can be done effectively is a complicated question, but potentially there are three ways to change the punishment probability: (a) increase the quantity and/or quality of resources available to the local criminal justice system; (b) increase the efficiency with which resources are used by the system; and (c) reduce the existing constraints which may hinder the effectiveness of the criminal justice system. In terms of equation (5):

$$\rho = \rho_{c/a}(E) \cdot \rho_a(P_0, K) ,$$

E - efficiency of judicial system,

P_0 - quantity of police officers,

K - capital utilized by law enforcement (weapons, communications, etc.),

$$\frac{\partial \rho}{\partial E} = \underset{(+)}{\rho_a(P_0, K)} \cdot \underset{(+)}{\frac{\partial \rho_{c/a}}{\partial E}} > 0 , \tag{9}$$

$$\frac{\partial \rho}{\partial P_0} = \underset{(+)}{\rho_{c/a}(E)} \cdot \underset{(+)}{\frac{\partial \rho_a}{\partial P_0}} > 0 , \tag{10}$$

$$\frac{\partial \rho}{\partial K} = \underset{(+)}{\rho_{c/a}(E)} \cdot \underset{(+)}{\frac{\partial \rho_a}{\partial P_0}} > 0 . \tag{11}$$

Note that each marginal effect is dependent upon the value taken by the other variable. In other words, there is a certain degree of interaction between the probability of being arrested, which depends upon law enforcement capabilities, and that of being convicted, which is dependent, in turn, upon the efficiency of the judicial environment. Property (10), for example, implies that adding more police would increase "ρ_a" and therefore the punishment probability, "ρ". But if the courts happened to be lenient, or over-crowded, "ρ" may remain unchanged, due to the offsetting reduction in "$\rho_{c/a}$".

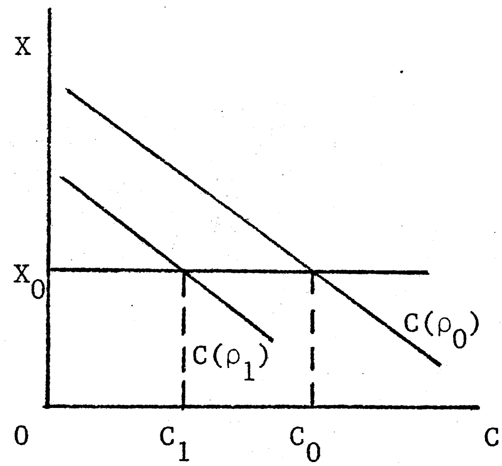

Figure 3. Crime Function

Next, consider the "technological" constraints facing the locality. The number of crimes prevented may be expressed in the following form:

$$C_p = C_p(\rho_{c/a}(E) \cdot \rho_a(P_0,K),\ X)\ , \tag{12}$$

C_p - Number of crimes prevented,

$$\frac{\partial C_p}{\partial P_0} = \underset{(+)}{\frac{\partial C_p}{\partial \rho}} \cdot \underset{(+)}{\frac{\partial \rho}{\partial P_0}} > 0\ , \tag{13}$$

$$\frac{\partial C_p}{\partial K} = \underset{(+)}{\frac{\partial C_p}{\partial \rho}} \cdot \underset{(+)}{\frac{\partial \rho}{\partial K}} > 0\ , \tag{14}$$

$$\frac{\partial C_p}{\partial E} = \underset{(+)}{\frac{\partial C_p}{\partial \rho}} \cdot \underset{(+)}{\frac{\partial \rho}{\partial E}} > 0\ . \tag{15}$$

Properties (13), (14), and (15) follow directly from (9), (10), and (11). Figures 4 and 5 depict the functional forms of Equations (12) and (13) respectively.

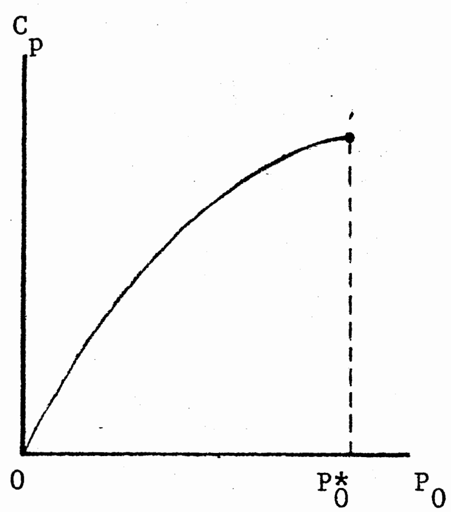

Figure 4. Crime Prevention Function

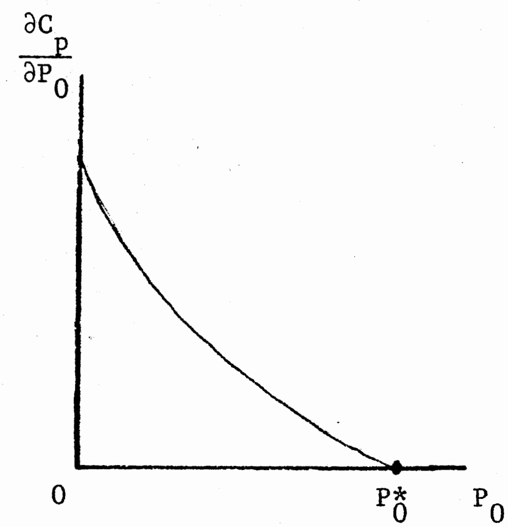

Figure 5. Marginal Product of Police Protection

Equation (13) may be thought of as the "marginal product" of police protection. It declines because of the usual reasons for diminishing marginal returns. Holding all other factors constant, including the available capital, each additional officer will prevent less and less criminal activity. At "P_0^*", the marginal product is zero. Changes in "E" and/or "K" will effect both "C_p" and "$\partial C_p/\partial P_0$", but perturbations in "X" (income distribution) will only effect the total of crime prevention ("C_p").

We may now incorporate these results in the urban crime function and analyze their implications (Figure 6).

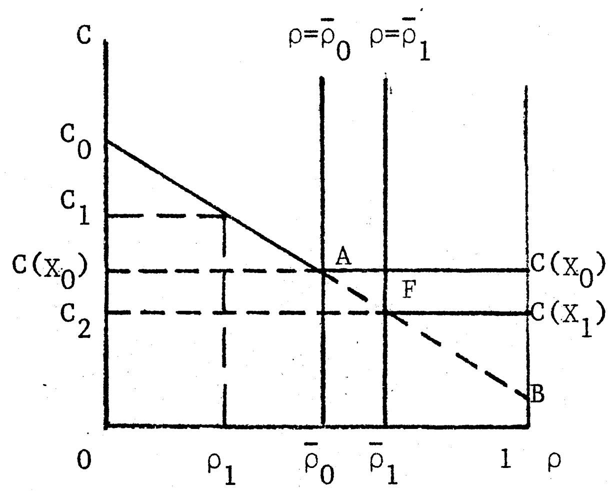

Figure 6. Crime and the Punishment Probability

The above diagram (Figure 6) is the relationship between crime and the punishment probability--drawn from the assumption of a given income distribution. Along segment "C_0A", crime may be reduced by increasing "ρ". (The crime prevented being, for example, $C_0 - C_1$.) This would be accomplished through increasing police protection ("P_0"), providing more capital ("K"), or altering the judicial environment ("E"). Moreover, if "ρ" were to remain fixed, movement to a more equitable distribution would also reduce crime (a shift downward in "$C_0AC(X_0)$"). Now, assume for expositionary purposes that "K" and "E" remain fixed along "C_0A", so that:

$$\rho = \rho_{c/a}(\bar{E}) \cdot \rho_a(P_0, \bar{K}),$$

and ρ will only vary when "P_0" changes. Thus, adding more police will increase "ρ", but will not reduce crime beyond "P_0^*" (page 27). This is represented by point "A", above, where:

$$\bar{\rho}_0 = \rho_{c/a}(\bar{E}) \cdot \rho_a(P_0^*, \bar{K}).$$

Increasing "ρ" past "$\bar{\rho}_0$" will not effect criminal activity, which is totally determined by the prevailing income distribution along "$AC(X_0)$". It follows, then, that crime may only be reduced in this region by either of two ways: redistributing income towards the poor ($C(X_0)$ to $C(X_1)$) and/or changing the availability of law enforcement capital or the effectiveness of the courts. Increasing the latter would reduce crime along the dashed segment "AFB". At point "F":

$$\bar{\rho}_1 = \rho_{c/a}(\bar{E}_1) \cdot \rho_a(P_0^*, \bar{K}_1). \qquad \bar{K}_1 > \bar{K}, \quad \bar{E}_1 > \bar{E}$$

To summarize, there are a number of policy options available to the local government. Assuming diminishing marginal returns to police protection, crime may be reduced through any of the aforementioned methods. In addition, redistributing income will also be effective. If, however, the marginal product of additional police is zero, then the only options involve redistribution, provision of crime protection equipment, or a relaxation of existing judicial constraints. Some statistical studies have suggested that the marginal product of police is indeed zero, which would, in reality, provide some justification for arguing that these options are real (Wilson, 1974). The way in which this study differs from previous analyses of crime is its specification of relative income comparisons as an important determinant of criminal behavior. A further implication of this approach, however, is the economic consequences associated with interdependent welfare comparisons. If people view the distribution of income from an individual perspective, independently of social rankings, then this, along with a condition of anomie, becomes a necessary condition for the committal of a criminal act. The degree of inequality is a public good with which some may be satisfied and others dissatisfied, but which everyone must "consume". To the extent that this is true, the market will not allocate correctly and the standard Pareto optimal outcomes will not apply. In the case of crime as a consumption externality associated with the distribution of income, the standard income-leisure choice will result in <u>more</u> criminal activity than is socially optimal. In the absence of voluntary redistributions, then, the role of the local government is to determine how to achieve this "optimal" level--whether through some kind of tax-subsidy scheme or increased public expenditures on deterrence. The remainder of this chapter will be devoted to a derivation of the optimal tax-subsidy and its implications.

A Condensed General Equilibrium Model

Condensations of the full general equilibrium model are common in the first-best literature, and will be utilized here to analyse consumer externalities. When analyzing consumer externalities, the detailed production relationships of the model are not really necessary, since the primary interest is in the interrelationships among consumers. Thus, the idea is to simplify the usual, full model by emphasizing the essential consumption elements.

In all public sector analysis, individuals' preferences are the fundamental demand data. For the urban area, ordinal preferences will be represented as follows (<u>without</u> externalities):

$$U_i = U_i(y_i, \delta_i), \quad i = 1, 2, \ldots, p \tag{16}$$

$$U_h = U_h(y_h, \delta_h), \quad h = 1, 2, \ldots, r$$

y_i, y_h - income of person i, h (alternatively, the "numeraire" good with $P_Y = 1$),

δ_i, δ_h - single factor supplied by person i, h.

Factor supplies enter the utility function with a negative sign. For example, if labor (L) is the only factor, then:

$$Le_i = 24 - L_i ,$$

$$Le_h = 24 - L_h ,$$

where 24 is the total hours in a day and 24 - L is a "good", leisure (Le).

Assuming production is efficient and can be represented as an implicit function:

$$F(Y, \Delta) \equiv 0 , \tag{17}$$

Y - aggregate income,

Δ - aggregate factor supply.

F( ) has the following property:

$$dF = \frac{\partial F}{\partial Y} dY + \frac{\partial F}{\partial \Delta} d\Delta \equiv 0 ,$$

$$\frac{dY}{d\Delta} = - \frac{\partial F/\partial \Delta}{\partial F/\partial Y} = MP_\Delta . \tag{18}$$

Thus, F( ) implicitly defines a production function whose derivative is simply the marginal product of the factor. The reciprocal of the "MP_Δ" is known as the "marginal factor requirement".

Market clearance requires that:

$$Y = \sum_{i=1}^{p} y_i + \sum_{h=1}^{r} y_h$$

and

$$\Delta = \sum_{i=1}^{p} \delta_i + \sum_{h=1}^{r} \delta_h .$$

Incorporating these into F( ):

$$F(Y = \sum_{i=1}^{p} y_i + \sum_{j=1}^{r} y_h, \ \Delta = \sum_{i=1}^{p} \delta_i + \sum_{h=1}^{r} \delta_h) \equiv 0$$

or

$$F(\sum_{j=1}^{n} y_i, \ \sum_{j=1}^{n} \delta_i) \equiv 0, \qquad p + r = n \tag{19}$$

In addition, producers do not care who receives (supplies) an additional unit of a good (factor), that is:

$$\frac{\partial F}{\partial Y} = \frac{\partial F}{\partial y_j} \quad , \qquad j = 1, 2, \ldots, n$$

$$\frac{\partial F}{\partial \Delta} = \frac{\partial F}{\partial \delta_j} \quad .$$

With the above behavioral equations and constraints, the social welfare maximization problem becomes simple to represent:

$$\text{Max:} \qquad W(\vec{U}_1, \vec{U}_2) \tag{20}$$

$$\underline{\text{s.t.:}} \qquad F(\sum_{j=1}^{n} y_j, \ \sum_{j=1}^{n} \delta_j) \equiv 0$$

$\vec{U}_1$ - vector of utilities of poor,

$\vec{U}_2$ - vector of utilities of rich.

The function W( ) is the usual Bergson (1938) social welfare function found in the literature. This model is general enough to generate the relevant Pareto conditions, whether accomplished through free choice or the urban government.

Setting up the Lagrangian:

$$Z = W(\vec{U}_1, \vec{U}_2) + \lambda F(\sum_{j=1}^{n} y_i, \sum_{i=1}^{n} \delta_i)$$

$$\frac{\partial Z}{\partial \delta_j} = \frac{\partial W}{\partial U_j}\frac{\partial U_j}{\partial \delta_j} + \lambda\frac{\partial F}{\partial \Delta} = 0 \qquad j, d = 1, 2, \ldots, n; \; j \neq d \tag{21}$$

$$\frac{\partial Z}{\partial \delta_d} = \frac{\partial W}{\partial U_d}\frac{\partial U_d}{\partial \delta_d} + \lambda\frac{\partial F}{\partial \Delta} = 0 \tag{22}$$

and

$$\frac{\partial W}{\partial U_j}\frac{\partial U_j}{\partial \delta_j} = -\lambda\frac{\partial F}{\partial \Delta} \qquad j = 1, 2, \ldots, n \tag{23}$$

$$\frac{\partial W}{\partial U_1}\frac{\partial U_1}{\partial \delta_1} = \frac{\partial W}{\partial U_2}\frac{\partial U_2}{\partial \delta_2} = \cdots = \frac{\partial W}{\partial U_n}\frac{\partial U_n}{\partial \delta_n} .$$

These are known as the interpersonal equity conditions, derived from considering a single factor ("δ") supplied by any two different people ($j \neq d$). Condition (23) says that interpersonal equity is achieved only if all factors are supplied such that social welfare is equalized for all, rich and poor, on the margin. This redistribution must be lump-sum, so that the dichotomy between equity and efficiency be maintained; in accordance with previous assumptions, the locality acts in an urban first-best environment.

The Pareto conditions can be derived by considering the first-order conditions with respect to one good (income) and one factor consumed by any one person:

$$\frac{\partial Z}{\partial y_j} = \frac{\partial W}{\partial U_j}\frac{\partial U_j}{\partial y_j} + \lambda\frac{\partial F}{\partial Y} = 0 \qquad j = 1, 2, \ldots, n \tag{24}$$

$$\frac{\partial Z}{\partial \delta_j} = \frac{\partial W}{\partial U_j} \frac{\partial U_j}{\partial \delta_j} + \lambda \frac{\partial F}{\partial \Delta} = 0 \tag{25}$$

$$\frac{(\partial W/\partial U_j)(\partial U_j/\partial \delta_j)}{(\partial W/\partial U_j)(\partial U_j/\partial y_j)} = \frac{\partial F/\partial \Delta}{\partial F/\partial Y} .$$

After cancelling the social welfare terms:

$$MRS^j_{\delta_j, y_j} = \frac{(\partial U_j/\ \partial \delta_j)}{(\partial U_j/\ \partial y_j)} = MP_\Delta \qquad j = 1, 2, \ldots, n \tag{26}$$

or, equivalently:

$$MRS^j_{y_j, \delta_j} = \frac{(\partial U_j/\partial y_j)}{(\partial U_j/\partial \delta_j)} = \frac{1}{MP_\Delta} , \qquad j = 1, 2, \ldots, n \tag{27}$$

where "$1/MP_\Delta$" is the marginal factor requirement. Under competitive conditions, the MRS will be the same for all individuals, and profit maximizing firms will hire factor "Δ" until $MP_\Delta = P_\Delta$ ($P_Y = 1$), so:

$$MRS^j_{y_j, \delta_j} = \frac{1}{P_\Delta} . \tag{28}$$

Condition (28) is the familiar optimal choice between income ("y_j") and leisure (if "δ" represented labor time).

If the urban area approximates a competitive market, then it will generate the Pareto-optimal conditions, but not necessarily the interpersonal equity conditions (23). If the population is not neutral with respect to the distribution of income, then the government must act according to the dictates of two additional sets of first-order conditions, the interpersonal equity conditions. With the existence of externalities, however, a perfectly competitive market will no longer generate a Pareto-optimal allocation of resources. Government intervention is required, not only to satisfy the interpersonal equity conditions, but to keep the urban economy on its first-best utility-possibilities frontier.

Externalities

To show in simple terms what effect externalities have on economic activity, consider the following example of two individuals, where one's consumption of a good, say "X", affects the other individual in an adverse manner (an external "diseconomy"):

$$U_1 = U_1(y_1, x_1, x_2) \; ,$$
$$U_2 = U_2\,(y_2, x_2) \; ,$$
$$\frac{\partial U_1}{\partial x_2} < 0 \; .$$

Since individual two does not take into account the effect he/she has on person one, the competitive market will yield the usual outcome:

$$MRS^{i}_{x_i y_i} = P_X = MC_X . \qquad i = 1,\, 2;\; P_Y = 1 \tag{29}$$

But this result is not Pareto optimal, because the external "cost" imposed on person one is not accounted for. The efficiency conditions (to be derived) are:

$$\begin{aligned} &MRS^{1}_{x_1 y_1} = MRS^{2}_{x_2 y_2} + MRS^{1}_{x_2 y_1} = P_X = MC_X, \qquad (P_Y = 1) \\ &MRS^{1}_{x_2 y_1} < 0. \end{aligned} \tag{30}$$

Note that in comparing individual two's competitive outcome (29) with condition (30), it is found that:

$$MRS^2_{x_2 y_2} > MRS^2_{x_2 y_2} + MRS^1_{x_2 y_1} ,$$

so the marginal cost of "X" is "too high", implying an over-abundance of resources devoted to the production of "X".

Handling criminal activity as a consumer externality requires additional assumptions regarding the behavior of the urban residents. It is assumed that all individuals, both rich and poor, view crime from an aggregate perspective; that is, no one cares who is actually committing a particular offense--individual identities are irrelevant. Thus, the external effect depends only upon the level of the prevailing income distribution. Moreover, as mentioned previously, the type of crime which appears in individuals' preference functions is victim-related (excluding homicides). In terms of equation (8) (page 25) the crime function may be expressed as:

$$C = \sum_{i=1}^{p} C_i = C(X(f(\vec{Y}_p), g(\vec{Y}_r))).$$

$\vec{Y}_p$ - vector of incomes of the poor

$\vec{Y}_r$ - vector of incomes of the rich

If both the rich and poor care only about the aggregate level of crime, then each will have a utility function of the following form:

$$U_i = U_i(y_i, \delta_i, C(X(f,g))), \tag{31}$$

$$U_h = U_h(y_h, \delta_h, C(X(f,g))),$$

$i = 1, 2, \ldots, p$

$h = 1, 2, \ldots, r$

"U_i" and "U_h" have the following properties:

$$\frac{\partial U_i}{\partial y_i} = \frac{\partial U_i}{\partial y_i} + \frac{\partial U_i}{\partial C}\frac{\partial C}{\partial X}\frac{\partial X}{\partial f} , \tag{32}$$

$$\frac{\partial U_h}{\partial y_h} = \frac{\partial U_h}{\partial y_h} + \frac{\partial U_h}{\partial C}\frac{\partial C}{\partial X}\frac{\partial X}{\partial g} ,$$

$$\frac{\partial U_i}{\partial y_s} = \frac{\partial U_i}{\partial C}\frac{\partial C}{\partial X}\frac{\partial X}{\partial f} > 0 , \qquad i,s = 1, 2, \ldots, p;\ i \neq s \tag{33}$$

$$\frac{\partial U_i}{\partial y_h} = \frac{\partial U_i}{\partial C}\frac{\partial C}{\partial X}\frac{\partial X}{\partial g} < 0 .$$

If the income of person "s" increases, the utility of individual "i" is affected because the relative position of the poor has improved (33), reducing aggregate criminal activity. Note that person "i" is not concerned with the identity of the individual (or individuals) who are better-off, but rather a changing distribution in "i's" favor. Alternatively, crime increases as the rich become more affluent, adversely affecting "i's" utility as the distribution shifts to group "g's" favor. When either person "i" or "h" increase their consumption (income), however, there are two different effects. The first term in (32) represents the independent effect on utility for both "i" and "h" as a result of the standard income-leisure choice. The second term is the indirect effect that each individual's choice has on the income distribution and therefore crime. In the case of the poor person, "i", he/she is aware that the distribution has changed favorably; this would reduce illegal activity and consequently increase "i's" utility. But since each individual reacts only to an aggregate formulation, in the process of the income-leisure choice, the indirect effect is ignored. This behavioral assumption is certainly reasonable; no one realizes that their own personal activity may affect them beneficially or otherwise. The same may be said for the typical rich person, or individual "h". In this case, the choice for greater income affects "h" adversely, through higher crime rates. As previously mentioned, this does not necessarily mean that the rich are the victims; rather, they fear the prospect of being victimized. Again, person "h" ignores the effect that his/her income choice has on the income distribution.

When an externality takes an aggregate form, it may be corrected by means of a single tax (or subsidy) placed upon each individual creating the external effect. This single tax is referred to in the literature as a "Pigouvian" tax, named for Pigou (1956). Other types of externalities are not so easily handled. For instance, external effects in which the identity of individuals matter, require that the government design a set of taxes, with a different tax levied on each individual. Given the problem of identification costs, correction of the externality may not be a Pareto-improvement. The above formulation of crime, with income and the prevailing income distribution being the externality-generating activity, lends itself to the aggregate analysis. Thus, from an intuitive standpoint, one might initially guess that the poor should be subsidized ($\partial U_i/\partial y_j > 0$) and the rich taxed ($\partial U_i/\partial y_h < 0$), since neither take into account the external effects while maximizing utility. This would reduce crime to the optimal level, as the income distribution became optimal from a social welfare standpoint.

The derivation of the single tax-subsidy case may be accomplished by considering social welfare maximization with preferences represented by (31):

Max: $W(\vec{U}_1, \vec{U}_2)$

s.t.: $F(\sum_{j=1}^{n} y_j, \sum_{j=1}^{n} \delta_i) \equiv 0 \,. \quad p + r = n$

As before, the interpersonal equity conditions with respect to any two individuals' supply of factor "δ" are:

$$\frac{\partial W}{\partial U_j} \frac{\partial U_j}{\partial \delta_j} = - \lambda \frac{\partial F}{\partial \Delta} \,, \qquad j = 1, 2, \ldots, n \tag{34}$$

the same result as in (23), page 34.

The Pareto-optimal conditions will be derived for each sub-group, "p" and "r". For the poor, the first order conditions with respect to the ith person's income and δ_i are:

$$\frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial y_i} + \sum_{j=1}^{n} \frac{\partial W}{\partial U_j} \frac{\partial U_j}{\partial C} \frac{\partial C}{\partial X} \frac{\partial X}{\partial f} = \lambda \frac{\partial F}{\partial Y} \qquad i = 1, 2, \ldots, p \tag{35}$$

$$\frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial \delta_i} = \lambda \frac{\partial F}{\partial \Delta} \,. \qquad i = 1, 2, \ldots, p \tag{36}$$

Equation (35) implies that individual "i" receives satisfaction from his/her own personal income, but this enjoyment also affects others, as well as himself/herself, through the distribution and criminal activity.

Dividing (35) by (36) yields the Pareto conditions:

$$\frac{\frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial y_i} + \sum_{j=1}^{n} \frac{\partial W}{\partial U_j} \frac{\partial U_j}{\partial C} \frac{\partial C}{\partial X} \frac{\partial X}{\partial f}}{\frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial \delta_i}} = \frac{1}{MP_\Delta} \,.$$

Since the interpersonal equity conditions have been satisfied (34), the social welfare terms $(\partial W/\partial U_j)$ cancel and:

$$\frac{\dfrac{\partial U_i}{\partial y_i}}{\dfrac{\partial U_i}{\partial \delta_i}} + \sum_{j=1}^{n} \frac{\dfrac{\partial U_j}{\partial C} \dfrac{\partial C}{\partial X} \dfrac{\partial X}{\partial f}}{\dfrac{\partial U_j}{\partial \delta_j}} = \frac{1}{MP_\Delta} ,$$

which may be written as:

$$MRS^{i}_{y_i \delta_i} + \frac{\partial X}{\partial f} \sum_{j=1}^{n} MRS^{j}_{X\delta_j} = \frac{1}{MP_\Delta} . \qquad i = 1, 2, \ldots, p. \tag{37}$$

Note that:

$$\frac{\partial U_j}{\partial X} = \frac{\partial U_j}{\partial C} \frac{\partial C}{\partial X} ,$$

so:

$$\frac{\dfrac{\partial U_j}{\partial X}}{\dfrac{\partial U_j}{\partial \delta_j}} = MRS^{j}_{X\delta_j} .$$

Utilizing the same procedure for the sub-group of the rich, the optimal condition is:

$$MRS^{h}_{y_h \delta_h} + \frac{\partial X}{\partial g} \sum_{j=1}^{n} MRS^{j}_{X\delta_j} = \frac{1}{MP_\Delta} . \qquad h = 1, 2, \ldots, r. \tag{38}$$

Both conditions (37) and (38) imply that the government should equate the personal rates of substitution between "Y" and "Δ", plus the total of all individuals' valuations of a small change in the income distribution in terms of δ_j.

At this point, a significant limitation of this approach must be noted. In deriving the Pareto-optimal conditions, it was assumed that the interpersonal equity conditions held, so the welfare terms cancelled. This ability to cancel, however, is not merely an arithmetic procedure.

It implies the ability of government to redistribute lump-sum in order to attain the "bliss" point. Without "first-best" policy tools, the terms will not cancel and conditions (37) and (38) will not be the necessary ones for a social welfare maximum. This is essentially why the assumption of a "first-best" environment was made at the beginning of this analysis. The derivation of the conditions with a non-optimal distribution are handled by a "second-best" approach--an analysis that this study will not consider.

Equations (37) and (38) may be rewritten as:

$$MRS^{i}_{y_i \delta_i} = \frac{1}{MP_\Delta} - \frac{\partial X}{\partial f} \sum_{j=1}^{n} MRS^{j}_{X\delta_j}, \tag{39}$$

$$MRS^{h}_{y_h \delta_h} = \frac{1}{MP_\Delta} - \frac{\partial X}{\partial g} \sum_{j=1}^{n} MRS^{j}_{X\delta_j}. \tag{40}$$

Thus, each poor and rich person's marginal rate of substitution differs from the marginal factor product by the same amount, the total of all external effects. Since good "Y" (or income) is the numeraire, $P_Y = 1$. Firms will set $P_\Delta = MP_\Delta$, or $1/P_\Delta = 1/MP_\Delta$, by profit maximization. Each individual, both rich and poor, will initially face the same price for supplying factor "Δ", assuming the marginal external effects are ignored. To achieve Pareto optimality, then, the government may place a tax on the rich, "t", and a subsidy on the poor, "s", equal to (see Appendix A):

$$t = -\frac{\partial X}{\partial g} \sum_{j=1}^{n} MRS^{j}_{X\delta_j}, \tag{41}$$

$$s = \frac{\partial X}{\partial f} \sum_{j=1}^{n} MRS^{j}_{X\delta_j}. \tag{42}$$

Alternatively, a proportional income tax and subsidy may be used, rather than a unit tax. As seen in Figure 7, with the subsidy, the poor will adjust so that the reciprocal of the price received by them for supplying the factor will be below the marginal factor requirement (or the price above the marginal product):

$$\frac{1}{P^p_\Delta} = \frac{1}{MP_\Delta} - s , \tag{43}$$

P^p_Δ - price received by poor.

The rich, after paying the tax, will receive a reciprocal price above the marginal factor requirement (or below the marginal product) (Figure 8):

$$\frac{1}{P^r_\Delta} = \frac{1}{MP_\Delta} + t , \tag{44}$$

P^r_Δ - price received by rich.

Both diagrams indicate an important property of the tax and subsidy: they must equal the sum of the external effects at the <u>optimal</u> level of the supplied factor. Setting them equal to the effect at the original competitive equilibrium is not correct ("ac" and "bd").

Although finding "t_{opt}" and "s_{opt}" is not a simple task, presumably the government can reach the correct formulation through a trial and error process. More specifically, if the functions "D", "$S_{private}$", and "S_{social}" characterize a competitive factor market and the relevant external effects, then a trial and error process will generate the optimal tax and subsidy in the limit (Baumol, 1972) (See Appendix B).

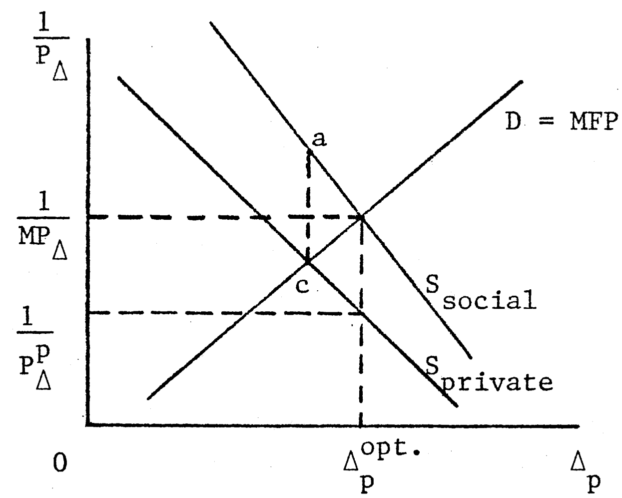

Figure 7. Labor Supply and Demand of the Poor

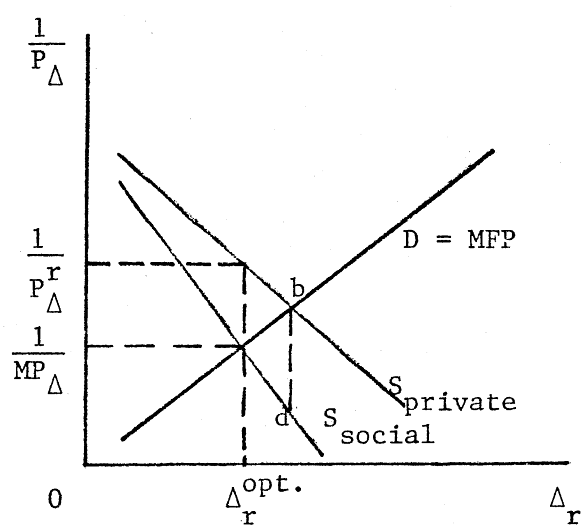

Figure 8. Labor Supply and Demand of the Rich

As a result of this tax-subsidy solution, the income distribution has changed in favor of the poor--a Pareto-optimal redistribution and criminal activity has also been reduced to its optimal level (Figure 9).

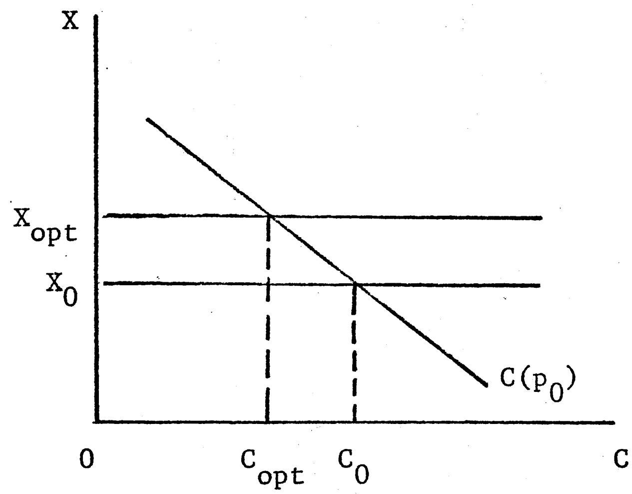

Figure 9. The Optimal Level of Crime

Voluntary Transfers

An obvious question is whether, in the absence of government intervention, charitable transfers would achieve the same result as the tax-subsidy approach. In other words, would a Pareto distribution prevail without a tax-transfer program, implying an optimum crime rate.

Consider the determination of an optimal transfer under the assumption that the costs of redistribution are shared equally by the rich. Each affluent individual has a utility function:

$$U_h = U_h(y_h,\ \delta_h,\ C(X(f,g)))\ , \qquad h = 1, 2, \ldots, r \tag{45}$$

whereas before:

$$f = f(\vec{Y}_p) ,$$

$$g = g(\vec{Y}_r) .$$

The tax paid by the rich is positively related to the dollar amount of income transferred:

$$T_h = T_h(R) \text{ , where} \tag{46}$$

R = income transferred to poor,

$$\frac{\partial T_h}{\partial R} > 0 .$$

Substituting (46) in (45) as a function of y_h, the optimal transfer occurs where:

$$\frac{\partial U_h}{\partial R} = \frac{\partial U_h}{\partial y_h} \frac{\partial y_h}{\partial T_h} \frac{\partial T_h}{\partial R} + \frac{\partial U_h}{\partial C} \frac{\partial C}{\partial X} \frac{\partial X}{\partial f} \sum_{i=1}^{p} \frac{\partial y_i}{\partial R} = 0 , \tag{47}$$

and:

$$\frac{\partial y_i}{\partial R} = 1; \qquad \frac{\partial y_h}{\partial T_h} = -1 .$$

If the tax burden is shared equally:

$$\frac{\partial T_h}{\partial R} = \frac{p}{r} \text{ , where}$$

p = number of poor,

r = number of rich.

Equation (47) now becomes:

$$\frac{\partial U_h}{\partial y_h} \left(\frac{p}{r} \right) = p \frac{\partial U_h}{\partial C} \frac{\partial C}{\partial X} \frac{\partial X}{\partial f} . \tag{48}$$

Thus, maximizing total utility from transferring income requires that the marginal benefit of the transfer (right side of (48)) equals the marginal cost (left side of (48)). The marginal benefit is the increase in utility of person "h" resulting from a reduction in crime, multiplied by the number of poor individuals. Note that a transfer initiated by any one rich person will benefit all affluent individuals. Therefore, a free-rider problem is likely to result without government intervention--there are not enough voluntary transfers. The marginal cost represents the individual's loss in utility from transferring income, multiplied by his/her share of the costs. According to (48), each rich person should prefer a government program to private charity, since all share the costs. Acting independently, the marginal costs to person "h" are $\partial U_h/\partial Y_h \cdot p$. It is easy to see that these costs would be reduced under general taxation ($p > p/r$).

In retrospect, there is no reason to expect that charitable behavior would be sufficient to achieve an optimal distribution and reduce crime significantly. Society must resort to the tax-subsidy scheme.

Policy

Up to this point, this analysis has considered one way of handling criminal activity--as a consumer externality. Preventing crime depends upon preserving or attaining a leptokurtic distribution of income. There is another option, however, which is increasing the probability of punishment. Referring to the diagram on page 26, increasing "ρ" will shift "C" to the left. Thus, the government may reach "C_{opt}" by a Pareto redistribution or by creating deterrents. The choice would depend upon the relative cost-benefit ratio associated with each option. In reality, though, the movement from a non-optimum to an optimum level of crime implies only a potential Pareto improvement, and not an actual Pareto improvement.

Consider first the costs involved with reducing crime through redistribution. The urban government may accomplish this through a general consumption tax on the rich and a consumption subsidy to the poor. Since this would correct the externality problem, there is no excess burden or efficiency cost involved. Note, however, that the subsidy is a multiple of the tax ((41) and (42)), or

$$s = -\frac{\partial X/\partial f}{\partial X/\partial g}\, t = \frac{dg}{df}\, t \qquad \frac{dg}{df} > 0 \ .$$

Thus, additional revenue must be raised in order to provide the subsidy. If this revenue is attained by distorting taxation, then the excess burden imposed would be a cost of redistribution. The locality could avoid this cost by lump sum taxation, and therefore face only the question of horizontal equity. So, it seems that the only real costs involved in this policy option would be administrative and compliance costs, and--perhaps the largest component--the costs of collecting information to determine the shape of the externality functions ($\partial X/\partial g \sum_{j=1}^{n} MRS^{j}_{X\delta_j}$ and $\partial X/\partial f \sum_{j=1}^{n} MRS^{j}_{X\delta_j}$) (see Appendix B).

The costs associated with reducing crime through deterrence are likely to be substantial. Assuming the marginal product of police protection to be zero, the alternatives involved would be to provide additional law enforcement capital and/or increase the effectiveness of the judicial system. Both options would entail opportunity costs, the loss of income in the private sector as resources are diverted into law enforcement. The types of costs here would be such things as communication devices, police vehicles, weapons, salaries of judges, administrative-legal fees, prison facilities, etc. The list is long, and all costs resulting from crime reduction should be included; that is, a decision to provide more enforcement capital will surely mean more arrests, thus the costs of providing prison facilities for those individuals must also be included. Moreover, financing these expenditures will result in a welfare cost if the method of taxation is distorting. Given that municipalities receive revenue from property and specific sales taxes, this burden is likely to be real and should be included in calculating the costs of deterrence.

Benefits of crime reduction would be the same for each policy option. Essentially, these would involve the reduction in costs that individuals bear as the result of crime, or the fear of it. For example, in the case of property crime, the benefits would be the value of the real property that was not destroyed as a result of less crime, and the reduction in psychic costs to victims and others, such as anger or fear. There are some "hidden" benefits, however, which are not readily apparent. If the property is not destroyed, then it is said that the net cost to society is zero, since only a redistribution among individuals takes place. But there is a real opportunity cost involved, which results from employment in crime-related activities. The value of the goods and services foregone because some people work as burglars, fences, etc., is a real cost. Less crime will reduce these costs and is a net benefit to urban residents.

Let us entertain the following hypothetical situation. Property crime is on the rise in New York City and the current administration is faced with the task of controlling it by the best means possible; that is, at minimum cost to taxpayers. Assume that the city government has all the "first-best" policy tools at its disposal--planners can simultaneously employ allocational policies and lump sum redistributions in order to maximize the social welfare of the urban population. Thus, the objective is to choose that option which has the highest benefit-cost ratio. In this case, there are only two options: redistribution or deterrence. Through some means of reaching a consensus, it has been determined that crime is to be reduced by some percentage. From an economic perspective, this may or may not be the optimal reduction, but given that the politicians are assumed to act in accordance with the preferences of the urban population, this reduction may, for all practical purposes, be considered optimal. The relationship between crime, the distribution of income, and the probability of punishment is:

$$C_{Rt} = \gamma_0 X_t^{\gamma_1} \rho_t^{\gamma_2} e^{\varepsilon_{1t}}, \qquad t = 1, 2, \ldots, T$$

C_{Rt} - urban crime rate,

γ_1, γ_2 - elasticity parameters,

X_t, ρ_t - as defined previously in text.

"T" is the current period and the parameters are assumed to be known. From this equation, the percentage changes needed in "X_t" and "γ_t" may be determined for any given reduction in crime:

$$\frac{\Delta X_t}{X_{t-1}} = \frac{\Delta C_{Rt}}{C_{Rt-1}} \cdot \frac{1}{\gamma_1},$$

$$\frac{\Delta \rho_t}{\rho_{t-1}} = \frac{\Delta C_{Rt}}{C_{Rt-1}} \cdot \frac{1}{\gamma_2}.$$

Since the additional benefits would be the same for each option, the selection criterion would reduce to a comparison of additional expenditures (costs). Two additional equations may be observed, relating the probability of punishment and the income distribution to expenditures on crime prevention and redistribution, respective:

$$\rho_t = \alpha_0 \, \xi_{1t}^{\alpha_1} \, e^{\varepsilon_{2t}},$$

$$X_t = \beta_0 \, \xi_{2t}^{\beta_1} \, e^{\varepsilon_{3t}} ,$$

ξ_{1t} - expenditures on crime prevention at time "t",

ξ_{2t} - expenditures on income redistribution at time "t".

Given "$\Delta X_t / X_{t-1}$" and "$\Delta\rho_t / \rho_{t-1}$", the required changes in expenditures are:

$$\Delta\xi_{1t} = \frac{\Delta C_{Rt}}{C_{Rt-1}} \cdot \frac{\xi_{1t-1}}{\gamma_2 \alpha_1} ,$$

$$\Delta\xi_{2t} = \frac{\Delta C_{Rt}}{C_{Rt-1}} \cdot \frac{\xi_{2t-1}}{\gamma_1 \beta_1} .$$

Therefore, the locality would choose to redistribute income as a course of action if:

$$\Delta\xi_{2t} < \Delta\xi_{1t}$$

or,

$$\frac{\Delta C_{Rt}}{C_{Rt-1}} \cdot \frac{\xi_{2t-1}}{\gamma_1 \beta_1} < \frac{\Delta C_{Rt}}{C_{Rt-1}} \cdot \frac{\xi_{1t-1}}{\gamma_2 \alpha_1} ,$$

$$\frac{\gamma_2}{\gamma_1} < \frac{\beta_1}{\alpha_1} \frac{\xi_{1t-1}}{\xi_{2t-1}} .$$

This expression depends upon the parameters of the three equations and the ratio of the previous period's expenditures. If $\gamma_1 > \gamma_2$ then to achieve any given decrease in urban crime requires a larger percentage increase in the probability of punishment than in the income distribution. This is not sufficient for accepting the redistribution option, though, because of the differences in expenditure elasticities and the expenditure weights. The smaller "γ_2" is relative to "γ_1", and the larger "$\beta_1 \xi_{1t-1}$" is relative to "$\alpha_1 \xi_{2t-1}$", the more the locality should choose the Pareto redistribution option.

Since the parameters of the three equations are not actually known, they must be estimated if comparisons are to be made between the additional costs of each option. If a time ordered sample were selected, the corresponding estimators of the parameters may be used in the comparison:

$$\frac{\hat{\gamma}_2}{\hat{\gamma}_1} < \frac{\hat{\beta}_1}{\hat{\alpha}_1} \frac{\xi_{2t-1}}{\xi_{1t-1}} ,$$

or rewritten in terms of magnitudes:

$$\hat{S}* = \xi_{2t-1} \; |\hat{\gamma}_2 \hat{\alpha}_1| \; - \xi_{1t-1} \; |\hat{\gamma}_1 \hat{\beta}_1| < 0.$$

The properties of the estimator "$\hat{S}*$" are difficult to derive. Although we should theoretically be able to observe different values of "$\hat{S}*$" for different observed samples, in the above formulation, "$\hat{S}*$" will also vary with "ξ_{2t-1}" and "ξ_{1t-1}". Thus, fixed values of "$|\hat{\gamma}_2 \hat{\beta}_1|$" and "$|\hat{\gamma}_1 \hat{\alpha}_1|$" will not yield a unique "$\hat{S}*$", but "t-1" estimates.

In order to overcome this problem, it will be assumed that $\xi_{1t-1} \approx \xi_{2t-1}$, or that the past levels of crime prevention and redistribution expenditures are approximately equal. Under this condition, the cost comparisons become:

$$\frac{\gamma_2}{\gamma_1} < \frac{\beta_1}{\alpha_1}\ ,\ (\text{for } \xi_{1t-1} \approx \xi_{2t-1})$$

and the new statistic "$\hat{S}$",

$$\hat{S} = |\hat{\gamma}_2\hat{\alpha}_1| - |\hat{\gamma}_1\hat{\beta}_1|$$

may be utilized to test the relative cost effectiveness of the two options. It may be seen that this assumption is equivalent to a comparison of percentage changes of additional expenditures:

$$\frac{\Delta\xi_{2t}}{\xi_{2t-1}} < \frac{\Delta\xi_{1t}}{\xi_{1t-1}}\ ,$$

$$\frac{\Delta\xi_{1t}}{\Delta\xi_{2t}} > \frac{\xi_{1t-1}}{\xi_{2t-1}}\ ,$$

$$\Delta\xi_{1t} > \Delta\xi_{2t}\ . \qquad \xi_{1t-1} \approx \xi_{2t-1}$$

That is, if the bases are approximately equal, a comparison of additional costs is the same as a comparison of percentage changes.

For the City of New York, the mean ratio of crime prevention expenditures to social welfare expenditures over the period 1930-1981 was .8862. Since this ratio is not significantly different from one, the assumption of approximately equal bases does not seem to be unwarranted.

In the scenario outlined above, the locality was assumed to be in a position to choose between the best option, redistribution or deterrence, depending upon which was the lower-cost alternative. The urban government may then proceed to provide the additional funds by raising whateven additional revenue is needed. In other words, the budget is variable. This is a necessary condition for achieving optimality, for if planners face budget constraints then the required expenditures will not be forthcoming and the optimal level of crime will not be achieved. It may be noted that this requirement of variable budgets coincides with the "first-best" assumptions made throughout this study. Government budgets constraints force "second-best" solutions (Tresch, 1981).

CHAPTER IV

MODEL SPECIFICATION

Hypothesis and Variable Description

The formulation developed at the end of Chapter III lends itself to an empirical test of the viability of income redistribution as a means of reducing criminal activity. The null and alternative hypotheses would be of the following form:

$$H_0: \quad |\gamma_2\alpha_1| - |\beta_1\gamma_1| \geq 0 \ , \tag{49}$$

$$H_1: \quad |\gamma_2\alpha_1| - |\beta_1\gamma_1| < 0 \ . \tag{50}$$

The estimator utilized for the purposes of this test is:

$$\hat{S} = |\hat{\gamma}_2\hat{\alpha}_1| - |\hat{\beta}_1\hat{\gamma}_1| \ ,$$

where:

$$\hat{S} \xrightarrow{L} N(|\gamma_2\alpha_1| - |\beta_1\gamma_1|, \ \sigma^2_{\hat{S}}) \ .$$

"$\hat{S}$" has a limiting normal distribution ($\xrightarrow{L}$) (see Appendix C), and H_0 will be rejected when:

$$\left(\frac{\hat{S}}{\hat{\sigma}_{\hat{S}}}\right)_{obs} < z_{c,\alpha} \ ,$$

where "$z_{c,\alpha}$" is the critical value of a standardized normal distribution.

The parameter estimates, "$\hat{\gamma}_2$", "$\hat{\alpha}_1$", "$\hat{\beta}_1$", and "$\hat{\gamma}_1$" will be obtained from the following three equations:

$$C_{Rt} = \gamma_0 \, X_{t-1}^{\gamma_1} \, \rho_{t-1}^{\gamma_2} \, e^{\varepsilon_{1t}} , \tag{51}$$

$$\rho_t = \alpha_0 \, \xi_{1t}^{\alpha_1} \, e^{\varepsilon_{2t}} , \tag{52}$$

$$X_t = \beta_0 \, \xi_{2t}^{\beta_1} \, e^{\varepsilon_{3t}} , \tag{53}$$

$$t = 1, 2, \ldots, 52 ,$$

where C_{Rt} - an "index" of property crime,

X_t - the distribution of income,

ρ_t - the probability of punishment,

ξ_{1t} - expenditures on crime prevention, and

ξ_{2t} - expenditures on social welfare programs.

The relevant population for this study is New York City, which encompasses the five boroughs of the Bronx, Brooklyn, Manhattan, Queens, and Staten Island.

It is assumed that the crimes under consideration, robbery, burglary, larceny, and motor-vehicle theft are committed largely by those with incomes in New York City below the mean. This assumption does not seem to be unrealistic, since "the poor, uneducated individual with minimal skills is more likely to commit property crimes than the person higher up on the socioeconomic ladder" (U.S. National Commission on Law Observance and Enforcement, 1971, p.126). The NYC Police Department (December, 1981) has data on relative percentages by crime areas in the city over the 10 year period, 1971-1980. During this period, 66.7 percent, 40.1 percent, 58.4 percent, and 52.1 percent of the average number of offenses in robbery, burglary, larceny, and motor-vehicle theft respectively, were reported in the areas of Bedford-Stuyvesant in Brooklyn, South-Bronx in the Bronx, Jamaica in Queens, Harlem in Upper Manhattan, Jackson Heights in Queens, and Coney Island in Brooklyn. According to the New York City Department of Social Services Economic and Social Statistics (1971-1980), these areas are also among the most economically-depressed in the city. This assumption of criminality among the poor is not meant to imply that lower income individuals are more "criminal" than the rest of the population--only that they are more likely to commit property-related crimes. To argue that the poor are more "criminal" than other classes on the basis of their offense rates for these crimes would be as indefensible as arguing that upper income groups are more "criminal" on the basis of their offense rates for white collar crimes.

The crime variable, C_{Rt}, is defined as an "index" of the crimes of robbery, burglary, larceny, and motor-vehicle theft derived from a canonical correlation analysis. This procedure is used to select the linear combinations of two sets of variables that maximize the correlations between the combinations (Morrison, 1976). The "index" is the linear compound of the crime variables which has the maximum correlation with the various compounds of "X_{t-1}" and "ρ_{t-1}". The full procedure is outlined in Appendix C.

In order for the variable X_t to have operational meaning, the arbitrary function, X( ), must be replaced with a specific one. One specification which has all the properties derived in Chapter III, is:

$$X_t = \frac{\bar{Y}_b}{\bar{Y}_a} \quad , \tag{54}$$

where $\bar{Y}_b$ - mean income of individuals below the mean, and

$\bar{Y}_a$ - mean income of individuals above the mean.

The definition of income used to calculate "$\bar{Y}_b$" and "$\bar{Y}_a$" is,

$$\text{Value Added} + \begin{cases} \text{City Public Assistance} \\ \text{AFDC Payments} \end{cases} - \begin{cases} \text{City Income Taxes} \\ \text{Federal Income Taxes} \end{cases}$$

and it is expressed in real terms.

It is assumed that the standard of living of those with incomes above the mean provide the basis for relative welfare comparisons of whose with incomes below the mean. "$\bar{Y}_a$" may then be considered an aggregate "reference group". In an affluent society, those at the bottom of the income distribution evaluate both their present and future economic position with reference to the incomes of those at the top. This assumption implies that the economic position of the more affluent city dwellers generates the general feelings of malevolence that result in criminal activity directed against the most available property or persons.

"ρ_t", the probability of punishment, was defined in Chapter III as:

$$\rho = \rho_{c/a} \cdot \rho_a$$

or, as the probability of being convicted given the event of being arrested and charged, times the probability of being arrested and charged. "ρ" may be rewritten as:

$$\rho = \frac{n(c \cap a)}{n(a)} \cdot \frac{n(a)}{n(C)} \quad ,$$

$n(c \cap a)$ - number of individuals arrested, charged, and convicted for committing property crime,

$n(a)$ - number of individuals arrested and charged for property crime, and

$n(C)$ - total number of property crime offenses.

As expressed above, "ρ" is an a posteriori probability, and is easily computed from the available data.

"ξ_{1t}" and "ξ_{2t}" are expenditures on crime prevention and redistribution respectively. These amounts were obtained from the adopted New York City budgets for the years 1930-1981. "ξ_{2t}", expenditures on social welfare, includes AFDC benefit payments, city benefit payments, and all administrative costs.

According to Equation (51), the level of criminal activity depends on the potential criminal's perception of the income distribution and the probability of punishment. These perceptions are not formulated instantaneously, but with a lag. "X" and "ρ" are assumed to be perceived with a one year lag (X_{t-1}, ρ_{t-1}).

A complete description of the above variables can be found in Appendix D.

Empirical Model

Equations (51), (52), and (53) may be rewritten as:

$$C^*_{Rt} = \gamma^*_0 + \gamma_1 X^*_{t-1} + \gamma_2 \rho^*_{t-1} + \varepsilon_{1t} \quad , \tag{56}$$

$$\rho^*_t = \alpha^*_0 + \alpha_1 \xi^*_{1t} + \varepsilon_{2t} \quad , \tag{57}$$

$$X^*_t = \beta^*_0 + \beta_1 \xi^*_{2t} + \varepsilon_{3t} \quad , \tag{58}$$

where the "*" represents the natural log of the variables. The endogenous and exogenous variables are C^*_{Rt}, X^*_t, ρ^*_t, and X^*_{t-1}, ρ^*_{t-1}, ξ^*_{1t}, ξ^*_{2t} respectively. The system of equations could be estimated individually using ordinary least squares, but more efficient (minimum variance) estimators may be obtained by allowing for the correlation of disturbances across equations (see Appendix C). The resulting estimators may then be used for testing the null hypothesis (49).

It is assumed that the disturbances, ε_{1t}, ε_{2t}, and ε_{3t} are correlated for different observations within equations and for the same and different observations across equations. This assumption may be tested by an examination of the estimated autocorrelation and cross-correlation function of the residual series, $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$, and $\hat{\varepsilon}_{3t}$. The null hypotheses of "white noise" would be:

$$H_0 1: E(\varepsilon_{it}\varepsilon_{jt-\tau}) = 0, \qquad (i,j) = (1,1)(2,2)(3,3), \quad \tau = 1, 2, \ldots, \ell$$

$$H_0 2: E(\varepsilon_{it}\varepsilon_{jt-\tau}) = 0, \qquad (i,j) = (1,2)(2,3)(1,3), \quad \tau = \pm 1, \pm 2, \ldots, \pm \ell$$

and the estimated functions:

$$r_\tau(\hat{\varepsilon}_i) = \sum_{t=\tau+1}^{T} \hat{\varepsilon}_{it}\,\hat{\varepsilon}_{jt-\tau} \Big/ \sum_{t=\tau+1}^{T} \hat{\varepsilon}_{it}^2, \qquad (i,j) = (1,1)(2,2)(3,3)$$

$$r_\tau(\hat{\varepsilon}_i\hat{\varepsilon}_j) = \sum_{t=1}^{T} \hat{\varepsilon}_{it}\,\hat{\varepsilon}_{jt-\tau} \Big/ \left(\sum_{t=1}^{T} \hat{\varepsilon}_{it}^2\right)^{\frac{1}{2}} \left(\sum_{t=1}^{T} \hat{\varepsilon}_{jt}^2\right)^{\frac{1}{2}}. \qquad (i,j) = (1,2)\ (2,3)\ (1,3)$$

It may be shown that under the null hypotheses (Tiao et al., 1980, p. 25):

$$r_\tau(\hat{\varepsilon}_i) \xrightarrow{L} N(0, T^{-1}),$$

$$r_\tau(\hat{\varepsilon}_i\hat{\varepsilon}_j) \xrightarrow{L} N(0, T^{-1}),$$

and

$$T \sum_{\tau=1}^{\ell} r_\tau^2(\hat{\varepsilon}_i) \xrightarrow{L} \chi_1^2(\ell),$$

$$T \sum_{\tau=-\ell}^{\ell} r_\tau^2(\hat{\varepsilon}_i \hat{\varepsilon}_j) \xrightarrow{L} \chi_2^2(2\ell) \ .$$

$H_0 1$ and $H_0 2$ would be rejected when $\chi^2_{1obs} > \chi^2_{1c,\alpha}$ and $\chi^2_{2obs} > \chi^2_{2c,\alpha}$ respectively. It is possible that one hypothesis would be accepted while the other is rejected. The estimation procedure is developed in Appendix C, under the assumption made above concerning the error structure, alternative specifications could be made if both hypotheses are not rejected.

The empirical results of the canonical correlation analysis, the estimation procedure, and the test of cost-effectiveness between redistribution and deterrence will be presented in Chapter V.

CHAPTER V

EMPIRICAL RESULTS

As mentioned in Chapter IV and derived in Appendix C, the crime "index", $C^*_{Rt} = Ln(C_{Rt})$, is obtained through a canonical correlation analysis. The procedure yielded only one statistically significant correlation, between the linear compounds "u_1" and "v_1", where:

$$u_1 = .8821 \cdot Ln(R) - .3463 \cdot Ln(B) + .3466 \cdot Ln(L) + .1320 \cdot Ln(MVT),$$

$$v_1 = .9119 \cdot Ln(X_{t-1}) + .1605 \cdot Ln(\rho_{t-1}) \ ,$$

$$\widehat{Corr}(u_1, v_1) = .7701^{**} \ ,$$

**significantly different from zero at $\alpha = .01$ ($P(F > F_{obs}) = .0001$),

R - Robbery,

B - Burglary,

L - Larceny,

MVT - Motor Vehicle Theft.

The correlation vector between "u_1" and the variables "Ln(R)", "Ln(B)", "Ln(L)", and "Ln(MVT)" is:

	Ln(R)	Ln(B)	Ln(L)	Ln(MVT)
u_1	(.9974	.9605	.9374	.9684) .

Since all of these correlations are high and positive, "u_1" is "capturing" virtually all the movement in property crime rates--high or low values of "u_1" are very likely to be associated with high or low values of robbery, burglary, larceny, and motor-vehicle theft. "u_1" is thus a "weighted-average" or "index" of property crime, representing the aggregate crime rate presented in Chapter III ($u_1 = \mathrm{Ln}(C_{Rt}) = C^*_{Rt}$).

In order to determine the structure of the variance-covariance matrix of the disturbances, the null hypotheses presented in the preceding chapter will be tested utilizing the residuals of the three-equation system:

$$H_0 1: E(\varepsilon_{it}\varepsilon_{jt-\tau}) = 0 \qquad (i,j) \qquad \tau = 1, 2, \ldots, 24$$

$$H_0 2: E(\varepsilon_{it}\varepsilon_{jt-\tau}) = 0 \qquad (i,j) \qquad \tau = \pm 1, \pm 2, \ldots, \pm 24$$

For $H_0 1$, the "χ^2" statistics are:

$$T \sum_{\tau=1}^{24} r_\tau^2(\hat{\varepsilon}_1) = 51 \cdot .7376 = 37.6213,$$

$$T \sum_{\tau=1}^{24} r_\tau^2(\hat{\varepsilon}_2) = 51 \cdot .7249 = 36.9684,$$

$$T \sum_{\tau=1}^{24} r_\tau^2(\hat{\varepsilon}_3) = 51 \cdot .7493 = 38.2158,$$

and for $H_0 2$:

$$T \sum_{\tau=-24}^{24} r_\tau^2(\hat{\varepsilon}_1\hat{\varepsilon}_2) = 51 \cdot 1.2425 = 63.3651,$$

$$T \sum_{\tau=-24}^{24} r_\tau^2(\hat{\varepsilon}_1\hat{\varepsilon}_3) = 51 \cdot 1.1807 = 60.2157,$$

$$T \sum_{\tau=-24}^{24} r_\tau^2(\hat{\varepsilon}_2\hat{\varepsilon}_3) = 51 \cdot 1.2665 = 64.589.$$

It may be seen from the above that the first hypothesis of no serial correlation within equations has been rejected, while the second hypothesis has been accepted, since $\chi^2_{1c,.05} = 36.42$ and $\chi^2_{2c,.05} = 65$. This implies a disturbance variance-covariance matrix of the form:

$$E(\varepsilon\varepsilon') = \begin{pmatrix} \phi_{11} & \sigma^{12}I & \sigma^{13}I \\ \sigma^{21}I & \phi_{22} & \sigma^{23}I \\ \sigma^{31}I & \sigma^{32}I & \phi_{33} \end{pmatrix} = \phi,$$

where:

$$\phi_{ij} = \begin{pmatrix} \sigma_0^{ij} & \sigma_1^{ij} & \cdots & \sigma_{T-1}^{ij} \\ & \ddots & \ddots & \sigma_1^{ij} \\ & & \ddots & \vdots \\ & & & \sigma_0^{ij} \end{pmatrix}, \quad (i,j) = \begin{matrix}(1,1)\\(2,2)\\(3,3)\end{matrix}$$

$$\sigma_\tau^{ij} = E(\varepsilon_{ij}\varepsilon_{jt-\tau}), \quad (i,j) \qquad \tau = 0, 1, 2, \ldots, T-1$$

and:

$$\sigma^{ij} = E(\varepsilon_{it}\varepsilon_{jt}) . \qquad i,j = 1, 2, 3, i = j.$$

The errors are correlated within equations for different observations and across equations for the same observations.

Knowledge of the restrictions on the error structure allows the application of a simple two-step estimation procedure to the three equations. First, each model is adjusted so as to eliminate the first order serial correlation within equations. This is accomplished by modeling the OLS residuals as a first-order autoregressive process and then transforming the system.

$$a_i(B)\hat{\varepsilon}_{it} = \hat{v}_{ij} \qquad a_i(B) = (1 - \hat{a}_i B) \qquad i = 1, 2, 3$$

$$a_1(B)C^*_{Rt} = \gamma^*_0 a_1(B) + \gamma_1 a_1(B)X^*_{t-1} + \gamma_2 a_1(B)\rho^*_{t-1} + \hat{v}_{1t} ,$$

$$a_2(B)\rho^*_t = \alpha^*_0 a_2(B) + \alpha_1 a_2(B)\xi^*_{1t} + \hat{v}_{2t} ,$$

$$a_3(B)X^*_t = \beta^*_0 a_3(B) + \beta_1 a_3(B)\xi^*_{2t} + \hat{v}_{3t} .$$

The second stage is to simultaneously estimate all parameters in the above system allowing for the cross-equation correlation of the disturbances (Zellner, 1963). The final estimates are:

$$\hat{C}^*_{Rt} = 4.3445 - 1.1125\, X^*_{t-1} - .5163\, \rho^*_{t-1} , \tag{59}$$
$$(5.889)^x \quad (-2.0055)^x \quad (-3.511)^x$$

$F_{obs} = 7.85$ $\quad R^2 = .2504$ $\quad a_1(B) = (1 - .1628\ B)$

$P(F > 7.85) = .0011$ $\quad DW = 1.8324$ $\quad \hat{\sigma}^{11}_0 = 1.558$

$$\hat{\rho}^*_t = -1.2166 + .1299\, \xi^*_{1t} , \tag{60}$$
$$(-8.1061)^x \quad (3.756)^x$$

$F_{obs} = 14.35$ $\quad R^2 = .2302$ $\quad a_2(B) = (1 - .1976\ B)$

$P(F > 14.35) = .0004$ $\quad DW = 1.8491$ $\quad \hat{\sigma}^{22}_0 = .0445$

$$\hat{X}^*_t = -2.3789 + .2406\, \xi^*_{2t} , \tag{61}$$
$$(-15.2001)^x \quad (7.6709)^x$$

$F_{obs} = 58.78$ $\quad R^2 = .5505$ $\quad a_3(B) = (1 - .1603\ B)$

$P(F > 58.78) = .0001$ $\quad DW = 1.8644$ $\quad \hat{\sigma}^{33}_0 = .0543$

( ) - "z" values in parentheses,

x - denotes significant at .05 level, and

DW - Durbin-Watson statistic.

All coefficients are highly significantly different from zero except that of the income distribution in equation (59). The probability of observing a value less than -2.0055 is approximately .023, while the critical region for the two-tailed test is .025. It is interesting to note that the percentage change in the crime index due to a one-percentage change in the income distribution (-1.1125) is twice as large as the respective coefficient of the probability of punishment (-.5163). Moreover, the same can be said of the elasticities of crime prevention and social welfare expenditures.

At first glance, one might note that the relationship,

$$\frac{\hat{\gamma}_2}{\hat{\gamma}_1} < \frac{\hat{\beta}_1}{\hat{\alpha}_1} = .4641 < 1.8522$$

holds for the estimated coefficients. The problem in a statistical sense, though, is to determine whether the difference, $|\hat{\gamma}_2\hat{\alpha}_1| - |\hat{\gamma}_1\hat{\beta}_1|$, is significantly less than zero to warrant rejection of the null hypothesis:

$$H_0\colon\ |\gamma_2\alpha_1| - |\gamma_1\beta_1| \geq 0 \quad ,$$

and acceptance of the proposition that income redistribution may be the more cost-effective way of combatting property crime.

The statistic, "$\hat{S}$", mentioned in the beginning of Chapter IV and developed in Appendix C is:

$$\hat{S} = |\hat{\gamma}_2\hat{\alpha}_1| - |\hat{\gamma}_1\hat{\beta}_1| = -.2006 \ .$$

The estimated standard error of "$\hat{S}$" (see Appendix C) is:

$$\hat{\sigma}_{\hat{S}} = .1207 \quad .$$

Since "$\hat{S}$" is asymptotically distributed as a normal variate,

$$\frac{\hat{S}}{\hat{\sigma}_{\hat{S}}} \xrightarrow{L} N(0,\ 1) \ ,$$

and,

$$\frac{\hat{S}}{\hat{\sigma}_{\hat{S}}} = \frac{-.2006}{.1207} = -1.6619 \quad .$$

The critical value of "$\hat{S}/\hat{\sigma}_{\hat{S}}$" is -1.6449 at the .05 level, so it appears that redistribution had been a less expensive alternative for the city of New York, without any realization of this purpose. The probability of observing a value of -1.6619 or less is approximately .048.

In summary, the estimated model has elasticity coefficients of the sign and magnitude hypothesized. Moreover, the magnitudes are large enough to warrant the labeling of redistribution as the lower-cost alternative.

The Need for Sensitivity Analysis

Although the previous analyses and empirical results have tended to support the viability of redistribution as a means of reducing crime, many of the computed values had probabilities of being observed that close to the rejection probability. Because of this closeness, an alternative specification of the error structure could have affected the magnitudes of the parameter estimates enough to reject the superior cost-effectiveness of the redistribution option. For example, the test applied on page 61 determined the absence or presence of serial correlation both within and across different equations. The first three statistics tested whether serial correlation existed within equations (56), (57), and (58), and the remaining three statistics tested the presence of serial correlation across the three equations (Chapter III). Note that the second and sixth observed values were very close to their respective critical values:

$$T \sum_{\tau=1}^{24} r_\tau^2(\hat{\varepsilon}_2) = 51 \cdot .7249 = 36.9684, \qquad \chi^2_{1c,.05} = 36.42,$$

$$T \sum_{\tau=-24}^{24} r_\tau^2(\hat{\varepsilon}_2\hat{\varepsilon}_3) = 51 \cdot 1.2665 = 64.589, \qquad \chi^2_{2c,.05} = 65.$$

If the null hypothesis $E(\varepsilon_{2t}\varepsilon_{2-\tau}) = 0$ had been accepted and the null hypothesis $E(\varepsilon_{2t}\varepsilon_{3t-\tau}) = 0$ rejected, then an alternative error structure would have resulted:

$$E(\varepsilon\varepsilon') = \begin{pmatrix} \phi_{11} & \sigma^{12}I & \sigma^{13}I \\ \sigma^{21}I & \sigma^{22}I & \phi_{23} \\ \sigma^{31}I & \phi_{32} & \phi_{33} \end{pmatrix} = \phi^*.$$

The only difference between this matrix and the one utilized in the actual estimation procedure are the partitioned matrices $\sigma^{22}I$ and ϕ_{32}, which represent no correlation across the disturbances in equation (57) and serial correlation between equations (57) and (58). The estimator of all regression parameters under the new specification is:

$$\hat{\Omega}^* = (Z' \phi^{*-1} Z)^{-1} Z' \phi^{*-1} Y,$$

and the original estimator was (see Appendix C):

$$\hat{\Omega} = (Z' \phi^{-1} Z)^{-1} Z' \phi^{-1} Y.$$

Thus different estimators will result, and the statistic, "$\hat{S}$", which was a function of these estimators, will also be sensitive to the new specification. Note, that the observed value of "$\hat{S}$", -.2006, was very close to its critical value of -.1985. Therefore, it is likely that an alternative disturbance structure may result in a different conclusion regarding which option has been more cost-effective.

CHAPTER VI

CONCLUSIONS

The objective of this study was the development of an alternative economic approach to urban property crime. A general equilibrium model was developed in which criminal activity was specified as an aggregate consumption externality associated with the distribution of income. When the more affluent urban residents make income-leisure choices, *ceteris paribus*, these decisions will adversely affect the welfare of others through a changing income distribution and thus crime. The income-leisure decisions of those on the lower end of the socio-economic scale, however, confer external benefits all urban residents, *ceteris paribus*, because crime is reduced through more *relative* equality. These external affects on welfare result because no resident takes into account the effects that changes in his/her own income has on the overall distribution. The typical "criminal" in this approach commits illegal acts partly because of a desire to improve relative income position. It is assumed that tastes and preferences are not exogenous, but are molded by our capitalist system which places great emphasis on monetary success with less regard for the means of achievement.

The empirical model developed and estimated for New York City supports the theoretical proposition that property crime results partly from relative economic deprivation. However, criminals also respond to changes in the criminal justice environment, as reflected in the probability of punishment. The model reflects the interaction among property crime, the probability of punishment, the income distribution, and public expenditures on crime prevention and social services. In testing the cost-effectiveness of redistribution versus deterrence as means of reducing crime, it was found that the hypothesis that deterrence is the lower-cost alternative could not be accepted, based upon the time-ordered sample.

This analysis and its empirical results have important implications for public policy. It is generally thought by economists that deterrence is the most effective means of reducing crime. This follows from the notion that certainty and severity of punishment will increase expected punishment costs, ceteris paribus. This relationship is considered invariant with respect to time or place; that is, the institutional setting is considered unimportant. The results of this study, however, indicate that one cannot simply analyze property crime outside the context of the social and economic structure. If individuals are not ultimately responsible for their own actions, then policy would clearly be mis-directed. The implication that more resources should be devoted to redistribution rather than deterrence, however, does not necessarily follow from this study. Equation (61) in Chapter V shows a .24 percent change in the income distribution with respect to a one percent change in social service expenditures. Over the period 1964-1981 administrative and salary costs for this city department grew 42.3 percent. Administering these programs more efficiently could yield a larger change in the income distribution for the same change in expenditures.

Finally, there are some implications for further research that could be explored. Alternative specifications for the income distribution could be utilized in the same model specification. For example, as mentioned in Appendix B, "X" could be of the form:

$$X_1 = \sum_{i=1}^{b} y_i \bigg/ \sum_{h=1}^{a} y_h ,$$

where b - number of individuals below aggregate mean, and

a - number of individuals above aggregate mean.

The function "X_1" is expressed as a ratio of total incomes. Another formulation is:

$$X_2 = \bar{Y}_{b,\alpha} \,/\, \bar{Y}_{a,\alpha} ,$$

where "α" is a specified percentile of the income distribution. For example, $\bar{Y}_{b,.25}$ would be the mean income of the lower 25 percent and $\bar{Y}_{a,.25}$ the mean income of the upper 25 percent. No matter what formulation of the income distribution is chosen, it is imperative that the function exhibit the theoretical properties derived in Chapter III. Any ratio would be consistent with these properties and would maintain the concept of "relative comparisons.

In addition to alternative income distribution specifications, different target populations may be utilized. The same analysis could be applied to a cross-sectional sample of cities in the United States, or to census-tracts of urban areas, given available data. New York City was chosen because of its compatibility with the theoretical model and because the necessary data were available to the author. Not only are there "pockets" of high crime areas in each borough, but these areas also tend to be associated with the lowest relative income levels (e.g., Harlem, South Bronx, Jamaica, Bedford-Stuyvesant). Given this "poverty amidst plenty", New York City proved to be an appropriate subject for this study.

It might be noted that using different populations and/or alternative measures of distribution with the same model specification as in this study, would enable a more thorough investigation of the relative income effects. Moreover, the hypothesis that redistribution is the lower cost alternative could be examined utilizing data that is less aggregated. This would be more useful from a policy perspective, since programs may then be "targeted" to those areas where relative income is the lowest.

BIBLIOGRAPHY

Baumol, W. J. "On Taxation and the Control of Externalities." American Economic Review, 51(1972):82-96.

Beccaria, Cesare. On Crimes and Punishments. New York: Bobbs-Merrill Co., 1963.

Becker, C. S. "Crime and Punishment: An Economic Approach." The Journal of Political Economy, 76(1968):169-193.

Bentham, J. The Works of Jeremy Bentham. Edited by John B. Bowrling. Edinburg: Singkin, Marshall, and Co., 1843.

Bergson, A. "A Reformulation of Certain Aspects of Welfare Economics." Quarterly Journal of Economics, 8(1938):111-126.

Block, M. J. and Heineke. "A Labor Theoretic Analysis of the Criminal Choice." American Economic Review, 65(1975):314-325.

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

Bonger, W. Criminality and Economic Conditions. Bloomington: Indiana University Press, 1969.

Chatfield, C. and Collins, A. J. Introduction to Multivariate Analysis. London: Chapman and Hall, 1980.

Cohen, Albert K. The Delinquency Subculture. Edited by Rose Giallombardo. New York: John Wiley and Sons, Inc., 1976.

Cohen, Lawrence E. "Robbery Victimization in the U.S.: An Analysis of a Nonrandom Event." Social Science Quarterly, 62(1981): 644-657.

Coljanni, Alfred C. Crime and the Social Condition. New York: McGraw Hill Book Co., 1959.

Cook, P. "The Correctional Carrot: Better Jobs for Parolees." Policy Analysis, 1(1975):11-15.

Easterlin, R. A. "Does Money Buy Happiness." The Public Interest, 10(1975):113-123.

Ehrlich, I. "A Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." Journal of Political Economy, 81, 3(1973):521-566.

Ferri, Enrico. Criminal Sociology. New York: Appleton and Co., 1896.

Fleischer, Belton. The Economics of Delinquency. Chicago: Quadrangle Books, 1966.

Forst, Brian. "An Analysis of Interrelationships Between Crime and Its Determinants." Washington, D.C.: Center for Naval Analysis, 1975. (Mimeographed.)

Friedlander, Stanley L. Unemployment in the Urban Core. New York: Praeger, 1972.

Glaser, Daniel and Rice, K. "Crime, Age, and Unemployment." American Sociological Review, 24(1959):679-686.

Gordon, David. "Class and the Economics of Crime." The Review of Radical Political Economics, 3(1971):51-75.

Hochman, Harold M. and Rodgers, James D. "Pareto Optimal Redistribution." American Economic Review, 49(1969):542-557.

Horton, Paul. "Problems in Understanding Motives." In The Economics of Crime and Punishment, ed. Simon Rottenberg. Washington, D.C.: American Enterprise Institute of Public Policy, 1975.

Kraus, M. and Mohring, H. "The Role of Polutee Taxes in Externality Problems." Economica, 81(1975):89-110.

Layard, P. R. G. and Walters, A. A. Micro-Economic Theory. New York: McGraw-Hill, Inc., 1978.

Leonard, J. W. "Dismissal for Off-The-Job Criminal Behavior." Monthly Labor Review, 90(1967):21-27.

Lenihaw, Kenneth J. Unlocking the Second Gate. Washington, D.C.: U. S. Department of Labor, Government Printing Office, 1976.

Lindeman, Richard H., Mercenda, Peter F., and Gold, Ruth Z. Introduction to Bivariate and Multivariate Analysis. Dallas, Texas: Scott, Foresman and Co., 1980.

Marx, Karl. The Marx-Engels Reader. Edited by Robert C. Tucker. Manifesto of the Communist Party. New York: W. W. Norton and Co., Inc., 1978.

Merton, Robert K. "Social Structure and Anomie." American Sociological Review, 3(1968):672-682.

Morrison, Donald F. Multivariate Statistical Methods. New York: McGraw Hill Book Co., 1976.

New York City Department of Social Services. Economic and Social Statistics. New York City, New York: 1930-1981.

New York City Police Department. Monthly Crime Reports. New York City, New York: January 1930-December 1981.

Office of the Mayor of the City of New York. New York City Budgets. New York City, New York: 1929-1981.

Pigou, A. C. The Economics of Welfare. 4th Edition. London: MacMillan, 1956.

Piore, Michael. "Public and Private Responsibilities in On-The-Job Training of Disadvantaged Workers." Cambridge, Mass.: Massachusetts Institute of Technology, 1978 (mimeographed).

Portnay, B. "Employment for Former Criminals." Cornell Law Review, 70(1970):121-130.

Pownal, George A. "Employment Problems of Released Prisoners." Manpower (January, 1971):26-31.

Rao, C. Radhakrismna. Linear Statistical Inference and Its Applications. New York: John Wiley and Sons, 1965.

Samuelson, Paul A. "The Pure Theory of Public Expenditures," Review of Economics and Statistics, 36(1954):387-389.

Staaf, Robert and Tannian, Francis, eds. Externalities, Theoretical Dimensions of Public Economy. New York: Dunellew Publishing Co., Inc.

Taggart, Robert. "Manpower for Criminal Offenders." Monthly Labor Review (August, 1972):17-25.

Thurow, Lester. "The Income Distribution as a Pure Public Good." Quarterly Journal of Economics, 85, 2(1971):92-102.

Tiao, G. C., Box, G. E. P., Grupe, M. R., Hudak, G. B., Bell, W. R., and Chang, I. The Wisconsin Multiple Time Series (WMTS-1) Program: A Preliminary Guide. Madison, Wisconsin: University of Wisconsin, 1981 (Mimeographed).

Tittle, Charles. "Punishment and Deterrence of Deviance." In The Economics of Crime and Punishment, Simon Rottenberg (ed.). Washington, D.C.: American Enterprise Institute of Public Policy, 1973.

Toby, Jackson. Affluence and Crime. Task Force Report on Juvenile Delinquency and Youth Crime. President's Commission on Law Enforcement and Administration of Justice. Washington, D.C.: Government Printing Office, 1967.

Tresch, Richard W. Public Finance, A Normative Theory. Plano, Texas: Business Publications, Inc., 1981.

Tulluck, Gordon. "Does Punishment Deter Crime?" The Public Interest, 36(1974):103-111.

U. S. Federal Bureau of Investigation. Uniform Crime Reports for the U. S. Washington, D.C., 1930-1981.

U. S. National Commission on Law Observance and Enforcement. Report on Causes of Crime. U. S. Department of Justice. Washington, D.C.: Government Printing Office, 1971.

Wilson, James W. "Do the Police Prevent Crimes?" The New York Times Magazine (October 12, 1974), pp. 18-19; 96-101.

Zellner, A. "Estimators of Seemingly Unrelated Regressions: Some Exact Finite Sample Results." Journal of American Statistical Association, 58 (1963):977-992.

APPENDICES

APPENDIX A

IMPOSITION OF A "RECIPROCAL" TAX

The imposition of a tax is not usually seen in the literature in terms of the reciprocal of the factor price. It will be shown that the two approaches are essentially equivalent and involve simple arithmetic manipulation.

Consider the factor involved to be labor time, measured negatively:

$$\delta_j = \ell_j < 0 . \quad j = 1, 2, \ldots, n$$

The individual problem is to make an optimal choice of y_j and ℓ_j:

Max: $\quad U_i = U_i(y_j, \ell_j)$

<u>s.t.</u> $\quad y_j = w\ell_j \qquad (P_y = 1)$

w - real wage.

The conditions for a maximum are:

$$MRS^j_{\ell_j y_j} = w$$

or:

$$MRS^j_{y_j \ell_j} = \frac{1}{w} .$$

In market equilibrium:

$$MRS^j_{y_j \ell_j} = \frac{1}{w} = \frac{1}{MP_L} .$$

The imposition of a proportionate income (consumption) tax drives a wedge between the wage received by suppliers and the wage received by producers:

$$w_S = w_P(1 - \tau),$$

w_S - wage to suppliers,

w_P - wage to producers,

τ - tax rate.

In equilibrium, after the tax:

$$w_P(1 - \tau) - w_P = t$$

or:

$$t = \tau w_P \tag{1}$$

where "t" is the unit tax.

In terms of the reciprocal wage, after the tax:

$$\frac{1}{w_P(1 - \tau)} - \frac{1}{w_P} = t$$

or:

$$t = \frac{\tau}{w_P(1 - \tau)} \quad . \tag{2}$$

Note, that in (1), "t" is negative, since $w_P > w_S$, and in (2), "t" is positive since $\frac{1}{w_P} < \frac{1}{w_S}$.

The divergence between "w_P" and "w_S" in terms of the reciprocal wage is:

$$\frac{1}{w_P(1 - \tau)} = \frac{1}{w_P} + \frac{\tau}{w_P(1 - \tau)}$$

and:

$$\frac{1}{w_P(1 - \delta)} = \frac{1}{w_P(1 - \delta)} \; .$$

This is equivalent to the usual case:

$$w_P(1 - \delta) = w_P(1 - \delta) \, ,$$

$$w_S = w_P - \delta w_P \; ,$$

or:

$$w_S = w_P - t \; .$$

APPENDIX B

DETERMINING THE OPTIMAL TAX AND SUBSIDY

The determination of the exact form of the externality functions may prove to be an impossible task. An approximate method will be offered here.

The largest problem facing the government would be determining $\sum_{j=1}^{n} MRS^{j}_{X\delta_j}$, the urban population's valuation of the income distribution in terms of a factor supplied. In its original fomulation, this would be impossible for two reasons. First, the costs involved of identifying each individual would be prohibitive. Second, given the non-exclusive nature of the good "X" (publicness), the government could not determine its true valuation because of non-revealed preferences. Therefore, a way around these problems would be dichotomize the population, assume valuations to be identical within each group, and then to estimate these group valuations.

$\sum_{j=1}^{n} MRS^{j}_{X\delta_j}$ may be written as:

$$\sum_{j=1}^{n} MRS^{j}_{X\delta_j} = \sum_{i=1}^{p} MRS^{i}_{X\delta_i} + \sum_{h=1}^{r} MRS^{h}_{X\delta_h} ,$$

p - number of poor,

r - number of rich.

Assuming that $MRS_{X\delta}$ are equivalent within each group, the above becomes:

$$\sum_{j=1}^{n} MRS^{j}_{X\delta_j} = pMRS^{p}_{X\delta} + rMRS^{r}_{X\delta} .$$

A rough estimate of $MRS^{p}_{X\delta}$ and $MRS^{r}_{X\delta}$ may be obtained from the linear models:

$$\delta_{tp} = \alpha_0 + \alpha_1 X_t + \varepsilon_t , \qquad (1)$$

$$t = 1, 2, \ldots, T$$

$$\delta_{tr} = \beta_0 + \beta_1 X_t + \mu_t . \qquad (2)$$

"δ_{tp}" and "δ_{tr}" could be hours of work effort in the ghetto district and other areas respectively, or the classification could be made in terms of income. The distribution function, which has all the properties of the hypothetical distribution function mentioned throughout this analysis, is:

$$X_t = \frac{\bar{Y}_p}{\bar{Y}_r} = \frac{r}{p} \frac{\sum_{i=1}^{p} y_i}{\sum_{h=1}^{r} y_h}$$

or

$$X_t = \frac{\sum_{i=1}^{p} y_i}{\sum_{h=1}^{r} y_h} \quad .$$

Thus, "$\hat{\alpha}_1$" and "$\hat{\beta}_1$" would provide rough approximations of $MRS^p_{X\delta}$ and $MRS^r_{X\delta}$.

Since $\partial X/\partial f = 1/\bar{Y}_r$ and $\partial X/\partial g = -(\bar{Y}_p/\bar{Y}^2_r)$, the externality functions become:

$$\frac{\partial X}{\partial f} \sum_{j=1}^{n} MRS^j_{X\delta_j} \approx \frac{1}{\bar{Y}_r} (p\hat{\alpha}_1 + r\hat{B}_1) \quad , \tag{3}$$

$$\frac{\partial X}{\partial g} \sum_{j=1}^{n} MRS^j_{X\delta_j} \approx -\frac{\bar{Y}_p}{\bar{Y}^2_r} (p\hat{\alpha}_1 + r\hat{B}_1) \, . \tag{4}$$

From these, the optimal tax and subsidy may be determined (see Figure 10).

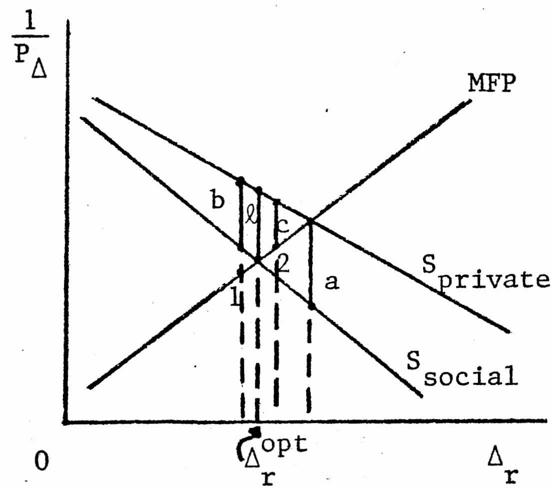

Figure 10. The Optimal Tax

The tax, "t", is initially set equal to "a", the marginal damages at the original equilibrium level. The new equilibrium after the tax will settle at "1" on MFP. The marginal damages have been reduced to "b", so the tax is adjusted to equal "b". This will bring the equilibrium level to "2". Again, the tax is adjusted to equal "c", and the trial and error process approaches "Δ_r^{opt}", with t = ℓ in the limit.

APPENDIX C

EMPIRICAL METHOD

Canonical Correlation

The objective is to derive an "index" of property crime through the procedure known as canonical correlation analysis (Lindeman, Merenda, Gold, 1980).

For the general case, assume that a vector of p + q random variables has been partitioned in the following manner:

$$\vec{Y}' = (\vec{X}_1' \ \vec{X}_2') \qquad \vec{X}_1' = (X_{11}, X_{12}, \ldots, X_{1p}),$$
$$\vec{X}_2' = (X_{21}, X_{22}, \ldots, X_{2q}).$$

The variance-covariance matrix of "$\vec{Y}$" may be partitioned as:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} .$$

From this population a sample of size n has been drawn, and the estimated variance-covariance matrix is:

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} .$$

The question is what are the linear combinations (canonical variates),

$$u_i = a_i'\vec{X}_1, \qquad i = 1, 2, \ldots, s$$
$$v_i = b_i'\vec{X}_2,$$

with the property that the sample correlation of u_1 and v_1 is greatest, the sample correlation of u_2 and v_2 is greatest among all linear compounds uncorrelated with u_1 and v_1, etc., for all s = min(p,q) possible pairs? It may be shown that the eigenvalues ("λ") obtained from the determinantal equations:

$$|S_{12}\ S_{22}^{-1}\ S'_{12} - \lambda S_{11}| = 0$$

or

$$|S'_{12}\ S_{11}^{-1}\ S_{12} - \lambda S_{22}| = 0$$

are the squared product-moment correlations of the iths linear compounds:

$$R^2_{u_i v_i} = \lambda_i \ .$$

To derive the weights, a_i and b_i, the following systems of linear equations may be solved:

$$(S_{12}\ S_{22}^{-1}\ S'_{12} - \lambda_i\ S_{11})a_i = 0 \ ,$$

$$(S'_{12}\ S_{11}^{-1}\ S_{12} - \lambda_i\ S_{22})b_i = 0$$

where λ_i is the ith largest root.

In order to interpret the results of a canonical correlation analysis, it is necessary to determine which variables in a set are most highly correlated with a given canonical variate and which least. The variance-covariance vectors of u_i, $\vec{X}_1$, and v_i, $\vec{X}_2$ are respectively:

$$\begin{aligned} \text{cov}(u_i, \vec{X}_1) &= E[(a'_i(\vec{X}_1 - \mu_1))((\vec{X}_1 - \mu_1)')] \\ &= a'_i\ \Sigma_{11} \ , \end{aligned}$$

$$\text{cov}(v_i, \vec{X}_2) = b'_i\ \Sigma_{22}$$

or, in sample form:

$$\widehat{\text{cov}}(u_i, \vec{X}_1) = a'_i\ S_{11} \ ,$$

$$\widehat{\text{cov}}(v_i, \vec{X}_2) = b'_i\ S_{22} \ .$$

Each of the elements of the above vectors may be transformed to correlation coefficients:

$$\widehat{\text{corr}}(u_i, X_{ij}) = \hat{\sigma}_{u_i X_{ij}} \Big/ \sqrt{a_i' \, S_{11} \, a_i} \cdot \sqrt{\hat{\sigma}^2_{X_{ij}}} \;, \qquad j = 1, 2, \ldots, p$$

$$\widehat{\text{corr}}(v_i, X_{ij}) = \hat{\sigma}_{v_i X_{ij}} \Big/ \sqrt{a_i' \, S_{22} \, a_i} \cdot \sqrt{\hat{\sigma}^2_{X_{ij}}} \;. \qquad j = 1, 2, \ldots, q$$

Suppose a single variable was needed for an analysis that would "capture" most of the movements in a set of related variables. This single variable should be positively and significantly related to its parent set of variables. Canonical correlation provides a method for selecting such an "index". Consider the crime equation formulated in the text:

$$\text{Ln}(C_{Rt}) = \text{Ln}\ \gamma_0 + \gamma_1\ \text{Ln}(X_{t-1}) + \gamma_2\ \text{Ln}(\rho_{t-1}) + \varepsilon_t$$

where "$\text{Ln}(C_{Rt})$" was defined as an "index" of property crime. This "index" may be written as:

$$\text{Ln}(C_{Rt}) = a_1\ \text{Ln}(R) + a_2\ \text{Ln}(B) + a_3\ \text{Ln}(L) + a_4\ \text{Ln}(MVT) \;,$$

R - robbery,

B - burglary,

L - larceny,

MVT - motor-vehicle theft,

(see Appendix D - Data Sources).

To derive "$\text{Ln}(C_{Rt})$", the weights, a_w, $w = 1, 2, 3, 4$, must be chosen in some optimal fashion. In terms of the procedure outlined above, the $p+q = 4 + 2 = 6$ variables are:

$$\vec{Y} = (\vec{X}_1' \; \vec{X}_2'),$$

$$\vec{X}_1' = (\mathrm{Ln}(R),\ \mathrm{Ln}(B),\ \mathrm{Ln}(L),\ \mathrm{Ln}(MVT)),$$

$$\vec{X}_2' = (\mathrm{Ln}(X_{t-1}),\ \mathrm{Ln}(\rho_{t-1})),$$

and the linear combinations:

$$u_i = a_i' \vec{X}_1 \;, \qquad u = 1,\ 2$$

$$v_i = b_i' \vec{X}_2 \;.$$

Since there are only two pairs of canonical variates ($s = \min(p,q)$), the "index", "$\mathrm{Ln}(C_{Rt})$", would be that value of u_i which has the <u>maximum</u> correlation with v_i (largest eigenvalue). Once chosen, the relationship between "$\mathrm{Ln}(C_{Rt})$" and the property crimes may be examined through the correlation vector. These correlations should be positive and high (by convention, greater than .25) (Chatfield and Collins, 1980).

Interpreting "$\mathrm{Ln}(C_{Rt})$" through the crime equation presents no difficulty. Under the hypothesis that $\gamma_1 < 0$, increases in the distribution function would reduce the index, and consequently reduce property crime.

Theoretical Model Specification

The estimation of the three equation model may be accomplished through Ordinary Least Squares (OLS), but more efficient estimates would be obtained by allowing for correlations among the error terms, both within the same equation for different observations and across equations for the same and different observations.

The three equations may be written as:

$$C^*_{Rt} = \gamma^*_0 + \gamma_1 X^*_{t-1} + \gamma_2 \rho^*_{t-1} + \varepsilon_{1t},$$

$$\rho^*_t = \alpha^*_0 + \alpha_1 \xi^*_{1t} + \varepsilon_{2t},$$

$$X^*_t = \beta^*_0 + \beta_1 \xi^*_{2t} + \varepsilon_{3t},$$

$$t = 1, 2, \ldots, T,$$

where "*" denotes the natural log of the variables. In matrix notation, the model becomes:

$$Y = Z\Omega + \varepsilon$$

$$\begin{pmatrix} C^* \\ \rho^* \\ X^* \end{pmatrix} = \begin{pmatrix} W^* & 0 & 0 \\ 0 & \xi^*_1 & 0 \\ 0 & 0 & \xi^*_2 \end{pmatrix} \begin{pmatrix} \gamma \\ \alpha \\ B \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

where Y = 3Tx1 vector,

Z = 3Tx7 matrix,

Ω = 7x1 vector, and

ε = 3Tx1 vector.

According to the aforementioned assumptions, there is serial correlation within equations and cross-equation correlation of the disturbances:

$$\phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix} = E(\varepsilon\varepsilon') \;,$$

where:

$$\phi_{ij} = \begin{pmatrix} \sigma_0^{ij} & \sigma_1^{ij} & \cdots & \sigma_{T-1}^{ij} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \sigma_1^{ij} \\ & & \ddots & \vdots \\ & & & \sigma_0^{ij} \end{pmatrix}, \qquad \begin{array}{l} i,j = 1, 2, 3 \\ i = j \end{array}$$

$$\sigma_\tau^{ij} = E(\varepsilon_{it}\varepsilon_{jt-\tau}), \qquad \tau = 0, 1, 2, \ldots, T-1$$

and:

$$\phi_{ij} = \begin{pmatrix} \sigma_0^{ij} & \sigma_{-1}^{ij} & \cdots & \sigma_{-(T-1)}^{ij} \\ \sigma_1^{ij} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_{-1}^{ij} \\ \sigma_{T-1}^{ij} & \cdots & \sigma_1^{ij} & \sigma_0^{ij} \end{pmatrix}, \qquad \begin{array}{l} i,j = 1, 2, 3 \\ i \neq j \end{array}$$

$$\sigma_\tau^{ij} = E(\varepsilon_{it}\varepsilon_{jt-\tau}), \qquad \tau = 0, \pm 1, \pm 2, \ldots, \pm T-1$$

The most efficient estimator of "Ω" is therefore:

$$\hat{\Omega} = (Z'\phi^{-1}Z)^{-1} Z'\phi^{-1}Y ,$$

and the variance-covariance matrix:

$$\text{Cov }(\hat{\Omega}) = (Z'\phi^{-1}Z)^{-1}.$$

In practice, the elements of "ϕ" must be estimated. These estimators are:

$$\hat{\sigma}_{\tau}^{ij} = \sum_{t=\tau+1}^{T} \hat{\varepsilon}_{it}\hat{\varepsilon}_{jt-\tau} / T, \qquad i = j;\ \tau = 0, 1, 2, \ldots, T-1$$

$$\hat{\sigma}_{\tau}^{ij} = \sum_{t=1}^{T} \hat{\varepsilon}_{it}\hat{\varepsilon}_{jt-\tau} / T. \qquad i \neq j;\ \tau = 0, \pm 1, \pm 2, \ldots, \pm T-1$$

The distribution of "$\hat{S}$" is derived by resorting to large sample theory. All of the least-squares estimators have a limiting normal distribution:

$$\sqrt{T}\,(\hat{\gamma}_i - \gamma_i) \xrightarrow{L} N(0, \lim T\, \sigma^2_{\hat{\gamma}_i}), \qquad i = 0, 1, 2$$

$$\sqrt{T}\,(\hat{\alpha}_i - \alpha_i) \xrightarrow{L} N(0, \lim T\, \sigma^2_{\hat{\alpha}_i}), \qquad i = 0, 1$$

$$\sqrt{T}\,(\hat{\beta}_i - \beta_i) \xrightarrow{L} N(0, \lim T\, \sigma^2_{\hat{\beta}_i}). \qquad i = 0, 1$$

Let $\hat{S} = f(\hat{\gamma}_1, \hat{\gamma}_2, \hat{\alpha}_1, \hat{\beta}_1) = \hat{\gamma}_2\, \hat{\alpha}_1 - \hat{\beta}_1\, \hat{\gamma}_1$, then:

$$\sqrt{T}\,((\hat{\gamma}_1\, \hat{\alpha}_1 - \hat{\beta}_1\, \hat{\gamma}_1) - (\gamma_1\, \alpha_1 - \beta_1\, \gamma_1)) \xrightarrow{L} N(0, \lim T\, \sigma^2_{\hat{S}}).$$

The limiting mean and variance of "$\hat{S}$" is derived by a Taylor expansion around the points $(\gamma_1, \gamma_2, \beta_1, \alpha_1)$:

$$\frac{\partial \hat{S}}{\partial \hat{\gamma}_2} = \hat{\alpha}_1\,, \qquad \frac{\partial \hat{S}}{\partial \hat{\alpha}_1} = \hat{\gamma}_2\,, \qquad -\frac{\partial \hat{S}}{\partial \hat{\beta}_1} = \hat{\gamma}_1\,, \qquad -\frac{\partial \hat{S}}{\partial \hat{\gamma}_1} = \hat{\beta}_1\,.$$

The series is:

$$\hat{S} \approx (\gamma_2\, \alpha_1 - \beta_1\, \gamma_1) + \alpha_1(\hat{\gamma}_2 - \gamma_2) + \gamma_2(\hat{\alpha}_1 - \alpha_1) -$$

$$\gamma_1(\hat{\beta}_1 - \beta_1) - \beta_1(\hat{\gamma}_1 - \gamma_1)\,.$$

The asymptotic mean is therefore:

$$\lim E(\hat{S}) \approx \gamma_2\,\alpha_1 - \beta_1\,\gamma_1.$$

The variance of "$\hat{S}$" is:

$$\sigma^2_{\hat{S}} \approx \alpha_1^2\,\sigma^2_{\hat{\gamma}_2} + \gamma_2^2\,\sigma^2_{\hat{\alpha}_1} + \gamma_1^2\,\sigma^2_{\hat{\beta}_1} + \beta_1^2\,\sigma^2_{\hat{\gamma}_1} + 2\alpha_1\,\gamma_2\,\sigma_{\hat{\gamma}_2\hat{\alpha}_1} -$$

$$2\alpha_1\,\gamma_1\,\sigma_{\hat{\gamma}_2\hat{\beta}_1} - 2\alpha_1\,\beta_1\,\sigma_{\hat{\gamma}_2\hat{\gamma}_1} - 2\gamma_2\,\gamma_1\,\sigma_{\hat{\alpha}_1\hat{\beta}_1} -$$

$$2\gamma_2\,\beta_1\,\sigma_{\hat{\alpha}_1\hat{\gamma}_1} - 2\gamma_1\,\beta_1\,\sigma_{\hat{\beta}_1\hat{\gamma}_1}\,.$$

the limiting variance-covariance matrix of the estimators, "$\hat{\Omega}$" is:

$$\lim \operatorname{cov}(\hat{\Omega}) = \frac{1}{T}\,\operatorname{plim}\,\left(\frac{Z\,\phi^{-1}\,Z}{T}\right)^{-1} = \frac{1}{T}\,Q^{-1}\,,$$

The elements of which are the individual limiting variances and covariances of the estimators. Since $\sigma^2_{\hat{S}}$ is a linear combination of a subset of these elements,

$$\lim \sigma^2_{\hat{S}} \approx \alpha_1^2 \lim \alpha^2_{\hat{\gamma}_2} + \gamma_2^2 \lim \sigma^2_{\hat{\alpha}_1} + \gamma_1^2 \lim \sigma^2_{\hat{\beta}_1} + \beta_1^2 \lim \sigma^2_{\hat{\gamma}_1} +$$

$$2\alpha_1\gamma_2 \lim \sigma_{\hat{\gamma}_2\hat{\alpha}_1} - 2\alpha_1\gamma_1 \lim \sigma_{\hat{\gamma}_2\hat{\beta}_1} - 2\alpha_1\beta_1 \lim \sigma_{\hat{\gamma}_2\hat{\gamma}_1} -$$

$$2\gamma_2\gamma_1 \lim \sigma_{\hat{\alpha}_2\hat{\beta}_1} - 2\gamma_2\beta_1 \lim \sigma_{\hat{\alpha}_1\hat{\gamma}_1} + 2\gamma_1\beta_1 \lim \sigma_{\hat{\beta}_1\hat{\gamma}_1},$$

which is the limiting variance of "$\hat{S}$".

The variance of "$\hat{S}$" may be consistently estimated by replacing each of the parameter and variances-covariances by their corresponding estimators:

$$\text{plim } \hat{\Omega} = \Omega \ ,$$

$$\hat{\text{cov}}(\hat{\Omega}) = (Z'\hat{\phi}^{-1}Z)^{-1} \ ,$$

$$\text{plim } \hat{\text{cov}}(\hat{\Omega}) = \frac{1}{T} \text{ plim } \left(\frac{Z'\hat{\phi}^{-1}Z}{T} \right)^{-1}$$

$$= \frac{1}{T} \text{ plim } \left(\frac{Z'\phi^{-1}Z}{T} \right)^{-1}$$

$$= \text{lim cov}(\hat{\Omega}) \ .$$

So:

$$\text{plim } \hat{\sigma}^2_{\hat{S}} = \sigma^2_{\hat{S}} \ .$$

Under the null hypothesis, $\gamma_1\alpha_1 - \beta_1\gamma_1 \geq 0$:

$$\frac{\hat{S}}{\sigma_{\hat{S}}} \xrightarrow{L} W \sim N(0,1) \ .$$

But since $\sigma_{\hat{S}}$ is not known, it must be proven that:

$$\frac{\hat{S}}{\hat{\sigma}_{\hat{S}}} \xrightarrow{L} W \sim N(0,1) \ .$$

Let $r = \dfrac{\sigma_{\hat{S}}}{\hat{\sigma}_{\hat{S}}}$, and $\text{plim } r = \dfrac{\sigma_{\hat{S}}}{\hat{\sigma}_{\hat{S}}} = 1$. Then:

$$\frac{\hat{S}}{\sigma_{\hat{S}}} \xrightarrow{L} W, \quad \frac{\sigma_{\hat{S}}}{\hat{\sigma}_{\hat{S}}} \xrightarrow{P} 1 \Rightarrow \frac{\hat{S}}{\sigma_{\hat{S}}} \cdot \frac{\sigma_{\hat{S}}}{\hat{\sigma}_{\hat{S}}} \xrightarrow{L} 1 \cdot W \ ,$$

By a limit theorem in large sample theory (Rao, 1965). Therefore:

$$\frac{\hat{S}}{\hat{\sigma}_{\hat{S}}} \xrightarrow{L} W \sim N(0,1) \ .$$

The null hypothesis would be rejected when:

$$P\left(\frac{\hat{S}}{\hat{\sigma}} < z_{obs}\right) < \alpha$$

for some specified "α" level.

APPENDIX D

DATA SOURCES

Data Sources

1. C_{Rt} - A canonical variate which is a linear combination of the following property crime rates:

   a. R - number of robberies per 100,000 population

   b. B - number of burglaries per 100,000 population

   c. L - number of larcenies per 100,000 population

   d. MVT - number of motor-vehicle thefts per 100,000 population

   For a complete definition of crimes included in these categories, see United States Federal Bureau of Investigation, Uniform Crime Reports for the U.S., annually. Data was obtained from New York City Police Department, Monthly Crime Reports, January 1930-December 1981.

   Ideally, one wants the number of crimes committed; in fact, the city crime reports contain only crimes reported to the police, and this is known to be an underestimate of the true crime rate. One can only assume that the ratio of reported crimes to actual crimes has been constant over time.

2. X_t - As defined in text

   Data on the income distribution for each year were obtained from the New York City Department of Social Services, Economic and Social Statistics, 1930-1981. The distributions were expressed in real terms through deflation by the Consumer Price Index. Income is defined here as money income + city and federal cash transfers - city and federal income taxes, for the period 1945-1981. In the earlier years, 1930-1944, only unadjusted money income was available.

   In order to find the mean above the mean and the mean below the mean from data initially given in ten income classes, the following functions were fitted for each year:

a. $N_i = f(Y_i)$

b. $\bar{Y}_i = g(N_i)$

N_i - total number of people who have incomes up to income Y_i

$\bar{Y}_i$ - cumulating mean of incomes up to income Y_i

$i = 1, 2, \ldots, 10$ income classes

From the aggregate data, the overall mean, "$\bar{Y}$", was known for each distribution. Substituting $\bar{Y}$ into equation (a), "n_b", the number of individuals with income up to the mean was derived. Then, by substituting "n_b" into equation (b), the mean of those with incomes up to the mean, "$\bar{Y}_b$", was derived. Since $n_a = n_T - n_b$ and $\bar{Y} = \dfrac{\bar{Y}_b n_b + \bar{Y}_a n_a}{n_T}$, $\bar{Y}_b$ was also derived.

3. $n(c \cap a)$, $n(a)$, $n(C)$ - as defined in text

Data were obtained from New York City Police Department, Monthly Crime Reports, January 1930-December 1981 and United States Federal Bureau of Investigation, Uniform Crime Reports for the U.S., annually.

4. ξ_{1t}, ξ_{2t} - as defined in text

Expenditures on crime prevention, "ξ_{1t}", included budget allocations for the following categories:

1) Police Department

a) Crime prevention and control

b) Investigation and apprehension

c) Emergency service

d) Support

e) Employee fringe benefits

f) Employee salary and wage adjustments

2) Criminal Court

   a) Executive management

   b) Judicial

   c) Fringe benefits

   d) Employee salary and wage adjustments

3) Department of Corrections

   a) Executive management

   b) Administrative and departmental services

   c) House of detention for women

   d) Male detention institutions and court detention pens

   e) Rehabilitation

   f) Employee fringe benefits

   g) Employee salary and wage adjustments

   h) Overtime pay

"ξ_{2t}" included budget expenditures for the following social service categories:

1) Department of Social Services

   a) Executive management

   b) Departmental services

   c) Public assistance (includes AFDC benefits)

   d) Employee fringe benefits

   e) Employee salary and wage adjustments

Data were obtained from the New York City budgets, 1929-1981.

2

VITA

Lonnie K. Stevans

Candidate for the Degree of

Doctor of Philosophy

Thesis: URBAN PROPERTY CRIME AND THE DISTRIBUTION OF INCOME

Major Field: Economics

Biographical:

Personal Data: Born in Huntington, New York, November 26, 1951, the son of Michael and Lucille Stevans.

Education: Graduated from Oklahoma Military Academy, Claremore, Oklahoma, in May, 1969; received Associate of Arts degree from Oklahoma Military Academy in 1971; received Bachelor of Science degree from Hofstra University in 1979; received Master of Science degree from Oklahoma State University in July, 1980; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1983.

Professional Organizations: Phi Kappa Phi, Beta Gamma Sigma, Omicron Delta Epsilon.