MSU

OVERDUE FINES ARE 25¢ PER DAY
PER ITEM

Return to book drop to remove
this checkout from your record.

# MINIMUM MANDATORY SENTENCES AND PLEA BARGAINING:

# AN ECONOMIC PERSPECTIVE

By

Thomas Paul Chester

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

1979

# ABSTRACT

## MINIMUM MANDATORY SENTENCES AND PLEA BARGAINING: AN ECONOMIC PERSPECTIVE

By

Thomas Paul Chester

In recent years, plea bargaining in the courts by the prosecutor with defendants has received increasing attention from the public, who primarily have disapproved, and from economists, who have attempted to demonstrate theoretically the rationale for bargaining and the impact of defendant characteristics and prosecutorial resources on the probability and outcome of bargaining. This dissertation develops a stochastic model of the prosecutor's and defendant's behavior, which, in addition to allowing analysis of bargaining, provides a theoretical basis for empirical investigation of the effect of the imposition of a minimum mandatory sentence on the probability of bargaining and the outcome of negotiations between prosecutor and defendant. In this model, the prosecutor and defendant are viewed as maximizing their respective expected utility functions, subject to a budgetary constraint. The utility functions of the prosecutor and defendant are modeled as being partially dependent on the length of sentence received by the defendant upon conviction. The length of sentence, given the conviction charge, is not known with certainty; it is thus incorporated into the model as a random variable. The imposition of a mandatory minimum sentence is viewed as truncating the cumulative probability distribution associated with this sentence variable. The optimal charges sought by prosecutor and defendant before and after imposition of the mandatory minimum sentence are calculated. In the dissertation it is shown how changes in the optimal charges resulting from imposition of the minimum mandatory sentence are directly related to the shift in the marginal rate of substitution functions, of the prosecutor and defendant, of charge for other goods. These changes are calculated for their effect on trial as well as on plea-bargained cases. Conceptually, there are four possible outcomes of the effect of the minimum mandatory sentence on the probability of bargaining and charges sought. The theory allows for the development of conditions that were used to discriminate among the four cases. Empirical tests were developed and data from the Michigan Criminal Court System were used to test the theory. It was found that imposition of the mandatory minimum sentence resulted in a decrease of the prosecutor's marginal rate of substitution of other goods for change resulting in the prosecutor being a less-aggressive bargainer and the defendant's marginal rate of substitution of other goods for charge increasing, resulting in the defendant being a more-aggressive bargainer; thus the minimum mandatory sentence for a given offense and defendant characteristics resulted in the prosecutor making larger plea-bargaining concessions. The net effect was that the probability of bargaining was not changed by the imposition of the minimum mandatory sentence, nor was the expected length of prison incarceration altered by the change in the law.

To my wife
Mary
and my sons
John and David

ACKNOWLEDGMENTS

First, I would like to express my gratitude to the chairman of my dissertation committee, Dr. Robert H. Rashe, for his understanding, time, and guidance. His insights and unfailing perseverance made this dissertation possible. His willingness to support my endeavors is a debt that can never be repaid. For his kindness I will be forever grateful. I would also like to thank Dr. Norman P. Obst, Dr. Stephen Martin, and Dr. Anthony Y. C. Koo, my dissertation commitee, for their suggestions and support, without which this would have been a lesser dissertation.

My thanks to Dr. Judith Lachman, Dr. Jan Palmer, and Dr. James B. Ramsey for their introduction to the topic of plea bargaining. I would also like to thank Dr. J. Forester and Dr. Gilbert Low of the Massachusetts Institute of Technology who, during my stays at MIT, gave helpful suggestions on methodology.

To my colleagues, Larry Marsh, Ron Tracy, Joseph Stone, Robert Brueska, Steve Hustead, Art Havener, Robert Ranger, and Harvey Ollis, who, during my doctoral years, through their stimulating conversation nurtured my desire to explore the unknown, I say thanks.

I would especially like to thank my wife Mary, whose very special support made the quest for the Ph.D. all worthwhile. Also, I would like to thank my two children, John and David, both of whom were born during my stay at MSU, for they bore the burden of my studies as much as I. To my beloved mother and father, Rita and Paul Chester; my mother-in-law and father-in-law, Ann and Edward Mieduch; my supportive brother, Richard, and sister-in-law, Julie; and my uncle, Reverend Thomas J. Cain, I give my thanks for their unswerving faith which supported me through the ups and downs of graduate school.

Finally, I would like to thank Susan Cooley, who has exquisitely typed this dissertation.

Responsibilities for remaining errors in this dissertation belong to the author.

## FOREWORD

It seems appropriate to include some thoughts now that this work is finished.

It is a wonderful feeling to have this dissertation completed and the degree it represents in hand. At the beginning this moment was THE sought-after goal. Between that time and the present, a great deal of learning, living, and loving has taken place so that THE goal became intertwined with many others.

Learning. During the years spent on this degree, the greatest education occurred outside of the classroom. The insights into oneself and human nature cannot be taught, textbook style, to have any real meaning. One has to experience.

Living. Completion of this goal can be compared to waking from a night's sleep. The night, like most, was filled with dreams, most good and some not. Now that the dawn has come, it will be time to pursue the good dreams, already having forgotten the others.

Loving. This was expressed in many ways by many people who crossed THE goal's path. It was most evident in the form of support: educational, professional, financial, and emotional. It was the most important element that allowed for this accomplishment.

It is a wonderful feeling to be finished. The learning, living, and loving will continue to take place along with the understanding that fulfillment of a goal is not only connected to ability; although ability must be present, it is bound to desire. This desire cannot be lost in intertwined goals. It has to be allowed to flourish to become a reality.

To reach such a tremendous goal has at times been a trial of patience and a study of flexibility. It has been an exhausting victory, but well worth the price. In the end it is a sweet accomplishment that can almost be tasted.

Now the new day will be filled with the setting and realization of other goals, so with a prayer of thanksgiving it's time to move on.

Mary

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

## LIST OF CHARTS

# INTRODUCTION

Plea bargaining between prosecutor and defendant has only recently become a matter of public concern. Until the seminal work by Newman[1] in 1966, no systematic attempt had been made to describe the plea bargaining process or the extent to which it is used. Recently several economists have examined the issue of plea bargaining within the framework of utility maximization.[2] Available resources relative to caseload, quality of evidence, status of defendant prior to trial, and so forth have been analyzed and their effects on the likelihood of plea bargaining assessed. Macro level studies have focused on the determination of the percentage of cases that will be resolved via bargaining, while micro analysis has focused on the forces which result in a particular case being selected for plea bargaining. This dissertation examines the effects of variations in risk, as defined in Appendix A, on the probability that bargaining will take place in a particular case. This is the first stochastic micro model of plea bargaining. Variations in risk that do not involve the probability of conviction are shown to affect the bargaining space of the prosecutor and of the defendant. The model incorporates most of the results obtained by earlier efforts and allows analysis of certain real world forces not included in the existing literature. It also provides an explanation for an anomaly in a previous work as to why increases in severity of sentencing need not result in a reduction of the probability of bargaining.[3]

The model will be used to demonstrate that inclusion of a mandatory sentence for a given class of crimes can result in a reduced charge against defendants. This hypothesis is tested using data from Michigan involving crimes in which a gun was used, before and after a law was passed in early 1977 setting a mandatory minimum sentence for such cases.[4]

In Chapter I the structure in which the defendant and prosecutor operate is described. In Chapter II major research efforts by others on plea bargaining are outlined. In Chapter III a model of the prosecutor is developed, in which the prosecutor is viewed as maximizing expected utility, subject to a budget constraint. In particular, changes in the bargaining space stemming from changes in the level of risk are examined. In Chapter IV an analogous theory for the defendant is developed and the two are integrated into a unified bargaining theory. Chapter V contains a summary of the sample data. In Chapter VI the hypothesis is tested with regression analysis, using data obtained from Michigan courts. Chapter VII presents the conclusions of this dissertation. In Appendix A the meaning of changes in risk is outlined and related theorems, from the existing literature, that are used in the main body of the dissertation are presented.

## Footnotes--Introduction

[1] Donald J. Newman, Conviction, The Determination of Guilt or Innocence Without Trial, Report of the American Bar Foundation's Survey of the Administration of Criminal Justice in the United States (Boston: Little, Brown and Co., 1966).

[2] William M. Landes, "An Economic Analysis of the Courts," Journal of Law and Economics 14 (April 1971): 61-108; Gary S. Becker, "Crime and Punishment: An Economic Approach," Journal of Political Economy 76 (March/April 1968): 196-217; William M. Landes, "Rules for an Optimal Bail System: An Economic Approach," The Journal of Legal Studies 2,1 (1973): 79-106; Judith A. Lachman, "An Economic Model of Plea Bargaining in the Criminal Court System" (Ph.D. dissertation, Michigan State University, 1975).

[3] Landes, "An Economic Analysis of the Courts."

[4] On January 1, 1977, Michigan House Bill 5073 became effective, which mandates at least two years in prison for any person convicted of a felony while in possession of a firearm. "Sec. 227b. (1) A person who carries or has in his possession a firearm at the time he commits or attempts to commit a felony, except the violation of section 227 or section 227a, is guilty of a felony, and shall be imprisoned for 2 years. Upon a second conviction under this section, the person shall be imprisoned for 5 years. Upon a third or subsequent conviction under this section, the person shall be imprisoned for 10 years. (2) The term of imprisonment prescribed by this section shall be in addition to the sentence imposed for the conviction of the felony or the attempt to commit the felony, and shall be served consecutively with and preceding any term of imprisonment imposed for the conviction of the felony or attempt to commit the felony. (3) The term of imprisonment under this section shall not be suspended. The person subject to the sentence mandated by this section shall not be eligible for parole or probation during the mandatory term imposed pursuant to subsection (1).

# CHAPTER I

# THE PROSECUTORIAL STRUCTURE: FRAMEWORK FOR BARGAINING

## 1.1 Introduction

The prosecuting attorney's office is faced with a basic economic question. It is the desire of the prosecuting attorney to see that those cases with sufficient evidence to support the information lodged in them be brought to court and resolved, and that the guilty receive appropriate sentences. The prosecutor's office has sufficient resources for active trial prosecution of only a small percentage of the cases in its files. In many densely populated areas the system would come to a halt if more than 10 percent of all defendants were sentenced or acquitted by the process of trial.[1] Because of this inability to resolve all cases by trial, the prosecuting attorney must rely heavily upon negotiated pleas.[2] The process by which cases are resolved by negotiated plea rather than trial is the focus of this study.[3] It will be shown that lack of funds is not the only reason for resolving a case through a negotiated plea, sometimes referred to as plea bargaining or trading out. The basic economic question is how to allocate scarce resources among competing ends.

## 1.2 Forms of Plea Negotiation

Negotiated pleas take several forms. The prosecuting attorney may allow the defendant to plead guilty to a crime not consistent with the facts. For example, in the state of Michigan the penalties for burglary at nighttime are more severe than those for daytime burglary. Thus, in exchange for the plea of guilty many defendants are allowed to plead guilty to daytime burglary when, in fact, the act occurred at night.[4] Another form of inducement is the practice of allowing the defendant to plead guilty to an included crime which is contained in the information, but which is not considered to be the most punishable illegal act committed. A thief who, during a robbery, physically beats the victim might be charged with assault rather than robbery and assault, in exchange for a guilty plea. If multiple counts of a particular crime are pending against the accused, he may be charged on only a single count of the illegal act or may be promised concurrent sentences in return for a guilty plea.

A more subtle form of plea negotiation does not involve the alteration of the charges pending. Among judges there exists wide variation in sentencing procedures. If it is known that a particular judge is understanding of certain acts, although illegal, and imposes sentence accordingly, the defendants charged with these acts benefit by being sentenced by this judge. In contrast, we have all heard references to "hanging" judges.[5] In densely populated jurisdictions such as Detroit, the judges rotate on the bench. The prosecuting attorney's office sometimes will agree to a continuance of the case until it is possible for it to be heard by the appropriate judge, with the understanding that the defendant will then plead guilty to the original charge.[6]

The techniques by which defense counsel can steer a client before a particular judge are varied. Counsel may request adjournment of the case to gather more information. This does not assure that the case will come up before a different judge, but scheduling conflicts may result in the trial being assigned to a different judge. Defense counsel can go directly to the clerk who schedules cases for the courts and request that the case be heard in front of the desired judge. Counsel may mention that his client wishes to plead guilty. The clerk often will be receptive because such information helps him to schedule work loads correctly. Or, at the arraignment in front of the presiding judge a plea of not guilty may be entered, with the result that the defendant is bound over for trial in front of a different judge. At the time of the trial a plea of guilty is entered.

Steering becomes important and, typically, is sought if the offense is probational.[7] The obvious goal is to plead guilty before a judge who believes in probation, whatever the other concessions of the prosecuting attorney. Other inducements are possible, such as a promise by the prosecuting attorney to recommend to the judge that the defendant be considered for probation.

In some circumstances it is not necessary for the defendant to enter into negotiations to receive a reduction in the charge. In Michigan, armed robbery carries with it a legislated mandatory minimum sentence with a provision for no early parole. In many such cases, the prosecutor will automatically reduce the charge in an attempt to avoid the harsh minimum sentence.[8] This action could be simple recognition by the prosecutor that certain people cannot be convicted of armed robbery with its mandatory minimum sentence. Even in cases in which the prosecutor brings the defendant to court charged with armed robbery, the judge might not allow the defendant to plead guilty.

Little formal structure exists in the bargaining proceeding in most localities. It may consist of some hurried words during arraignment or reading of the docket. In most cases it is done in a hurried manner and in less than formal surroundings. If the defendant does not have counsel, the defendant usually is told on what charge the prosecutor's office will accept a guilty plea and asked if he will agree. Unless the defendant believes he can prove himself not guilty of the original charge, he typically will accept. If the defendant has obtained counsel, or counsel has been appointed for him, then counsel will negotiate with the prosecutor. This normally is not a complicated procedure. Counsel will present some facts about the defendant which, it is hoped, will lead to the acceptance of a reduced charge. This would include facts about employment, marital status, and length of time in the area, among other things. The greatest bargaining asset of counsel is precedent. Without defense counsel the accused usually has no way of knowing the going rate for a particular crime. The prosecutor is not obliged to tell. In a quickly held meeting, a compromise is attempted. These meetings occur in hallways and even in judges' chambers. Most lawyers secured by the poor are not known as trial lawyers; they tend to seek a settlement with the prosecutor. The agreed-upon charge is presented to the defendant, who normally accepts. After all, it is his counsel advising him to do so, "for his own good." If a compromise cannot be reached, a continuance can be granted. During this time defense counsel or the family of the defendant can acquire additional information with which to bargain.

A pretrial report gathered for use by the prosecutor and the judge frequently is not used due to staff limitations. These reports, when obtained, are not lengthy. They attempt to determine the character and position of the defendant in the community prior to the offense committed. This information can be used by the prosecutor in determining how to charge the defendant and by the judge in sentencing. Usually, however, the only information available to the prosecutor is the report of the police, which deals exclusively with the facts of the crime and the evidence to support the information presented. The information the prosecuting attorney acquires concerning the defendant is usually limited; even when an attempt is made to gather facts they are sparse.

## 1.3 Goals of the Prosecutor

It has been suggested[9] that the prosecutor can be viewed as if he had one or several of the following goals. First, his goal may be one of simple efficiency--disposing of each case as quickly as possible. Second, he might attempt to maximize the number of convictions and the severity of sentences. "In this role, the prosecutor must estimate the sentence that seems likely after a conviction at trial, and balance the 'discounted trial sentence' against the sentence he can insure through a plea agreement."[10] Third, the prosecutor may view himself as a judge attempting to do the right thing for the defendant. Finally, he may act as a legislator--reducing charges because the law is too harsh and unfair.[11]

Reduction in charges is often consistent with the goals of the prosecuting attorney. For example, in the case of the armed nighttime robber who is allowed to plead guilty to unarmed daytime robbery, the parole board will ask the convict to describe the type of gun used in the crime during its questioning of him. The answer to this question is important, since in cases involving guilty pleas the judge is concerned only with determining that the plea is voluntary and that the defendant understands the consequences of his plea. The information at hand from the court appearance will be minimal. Thus, it is hoped, that parole board will determine the real nature of the crime and be inclined to fit the punishment to the real act rather than the offense of record. The expected prison term may not vary significantly between conviction for the actual crime and conviction for the crime to which the defendant is finally allowed to plead guilty. The prosecutor has obtained a speedy conviction and the convict still receives the sentence given for armed robbery. Also, in setting the minimum sentence, the judge frequently takes into account the fact that this is a negotiated plea. (The maximum sentence usually is set by legislative act.)[12]

## 1.4 Strategies of the Prosecutor

In an attempt to improve his bargaining position, the prosecutor often charges the defendant with as many individual crimes as possible. This will include the most serious crime committed and all included crimes.[13]

This inflation in charges is basically of two types: horizontal and vertical. Horizontal overcharging is the multiplying of accusations against the defendant. The defendant may be charged with a separate offense for every criminal act in which he participated, or the single offense may be broken down into numerous component charges. Vertical overcharging is often used as a means of gaining leverage. In this case a single offense is charged, but at a higher level than circumstances seem to warrant. Thus the prosecutor might charge the defendant with grand theft when the facts seem to warrant petty theft.

The practice of overcharging appears to increase the leverage of the prosecutor in two ways. First, the maximum sentence that the defendant would receive is increased as the number or severity of the charges is expanded. The defendant would be less certain of his chances of being found not guilty of all charges. Second, the reading of a long indictment is alleged to have an effect upon the jury.[14] These two effects reinforce each other.

## 1.5 Goals of the Defendant

There exists debate as to what interests the defendant the most in this type of bargain--the possible reduction in maximum sentence or reduction of the minimum sentence. It is widely held that the defendant is most concerned about the maximum sentence. The risk exists that the maximum time might actually have to be served. Accordingly, the defendant is presumed willing to trade to get a reduction in its value. It is also argued that setting of the minimum is more important since it tends to be a more reliable measure of how much time will be served in jail.[15] This dissertation does not consider the effect of variations in the maximum upon the bargaining posture of the defendant. Variations in the minimum sentence, however, are analyzed and a means for determining the impact of variations in the minimum sentence on bargaining is developed.

## Footnotes--Chapter I

[1]This is a readily acknowledged fact by the courts. In the words of Warren Burger, "It is an elementary fact, historically and statistically, that the system of courts--number of judges, prosecutors and of court rooms--has been based on the premise that approximately 90 percent of all defendants will plead guilty, leaving 10 percent, more or less, to be tried." Warren Burger, "Address at the American Bar Association Annual Convention," New York Times, August 11, 1970, p. 1.

[2]Recently, analysis has been concluded which supports the contention that plea negotiation can be properly viewed as a "clearing house" mechanism. See W. Rhodes, "The Economics of Criminal Courts: A Theoretical and Empirical Investigation," The Journal of Legal Studies 5 (June 1976): 331-40.

[3]This is not the first study to analyze the plea bargaining process. For example, see J. Lachman, "An Economic Model of Plea Bargaining in the Criminal Court System" (Ph.D. dissertation, Michigan State University, 1975). This is a nonstochastic bargaining model of the pleas process focusing primarily on the concept of relative power. Recently M. Block and Robert Lind, "An Economic Analysis of Crimes Punishable by Imprisonment," The Journal of Legal Studies 4 (June 1975): 479-92, hinted at a stochastic approach to analysis of plea bargaining using the technique of unconstrained utility maximization.

[4]Frank J. Remington et al., Criminal Justice Administration (Indianapolis: Bobbs-Merrill Co., 1969). See chapter 5, "Charging--The Decision to Prosecute," for several such examples.

[5]Donald J. Newman, Conviction, The Determination of Guilt or Innocence Without Trial, Report of the American Bar Foundation's Survey of Administration of Criminal Justice in the United States (Boston: Little, Brown and Co., 1966), p. 122. "Two slum-area Negro defendants in Detroit were arrested for statutory rape involving three juvenile victims, also Negro. The charge was downgraded to disorderly conduct, and an assistant prosecutor explained: 'This kind of behavior is so prevalent among these people that we wouldn't have enough courts or jails if they were all charged and convicted. They don't see anything wrong with this behavior; all their friends are doing the same thing.'"

[6]The President's Commission on Law Enforcement and Administration of Justice, Task Force on Administration of Justice, Task Force Report: The Courts (Washington, D.C.: Government Printing Office, 1967), p. 23. "Other illustrations of disparity may be found in the results of the workshop sessions at the Federal Institute on

Disparity of Sentences. The judges were given sets of facts for several offenses and offenders and were asked what sentences they would have imposed. One case involved a 51-year-old man with no criminal record who pleaded guilty to evading $4,945 in taxes. At the time of his conviction he had a net worth in excess of $200,000 and had paid the full principal and interest on the taxes owed to the Government. Of the 54 judges who responded, 3 judges voted for a fine only; 23 judges voted for probation (some with a fine); 28 judges voted for prison terms ranging from less than 1 year to 5 years (some with a fine). In a bank robbery case the sentences ranged from probation to prison terms of from 5 to 20 years."

[7]Newman, p. 212.

[8]Dominick Vetri, "Guilty Pleas Bargaining: Compromises by Prosecutors to Secure Guilty Pleas," University of Pennsylvania Law Review 112 (1964): 865-68. Two hundred five chief prosecuting officers in the most populous counties of 43 states were questioned by Vetri as to reasons for plea bargaining. The responses of 53 prosecutors answering this question are tabulated below:

| Factor | Number of Prosecutors Reporting Factor as Influencing Plea Bargaining |
|---|---|
| a. Particular crime charged | 39 |
| b. Defendant's prior record | 43 |
| c. Court's ability to adequately cope with negotiated charge | 45 |
| d. Work load | 22 |
| e. Penalties of law too harsh | 18 |

See also: Albert W. Alschuler, "The Prosecutor's Role in Plea Bargaining," University of Chicago Law Review 36,50 (1968): 54.

[9]Ibid., pp. 52-53.

[10]William Landes, "An Economic Analysis of the Courts," Journal of Law and Economics 14 (April 1971): 61-108. Landes was the first to theoretically analyze the prosecutor's behavior using this role of the prosecutor as the basis for his model.

[11]George F. Cole, The American System of Criminal Justice (North Scituate: Duxbury Press, 1975). See p. 229 for roles of the prosecutor.

[12]Alschuler, pp. 97-98. Alschuler points out that prosecutors commonly maintain that a reduction in the level of a charge seldom affects the defendant's sentence. Alschuler goes on to quote a particular Detroit prosecutor who is primarily responsible for all negotiations within Degroit: "By reducing a charge, everyone seems to be happy. The defendant and his attorney feel that they have

accomplished something by being able to plead to an offense less severe than that for which there was actual guilt. . . , society is happy because the defendant, if sentenced, is put away for the same amount of time as if he had pleaded to the original charge."

[13] Ibid., pp. 85-90.

[14] Ibid., pp. 98-99. "The psychological effect of overcharging may be even more significant than its economic effect, for a trial commonly begins with a reading of the accusations to the jury. 'A long inditement enhances a weak case,' observes Boston defense attorney Paul T. Smith. 'The jurors may simply be overwhelmed by the magnitude of the prosecutor's allegations.' Los Angeles attorney Al Matthews adds, 'When a jury hears an endless list of charges, they tend to think: The District Attorney could be wrong once, but no one could be wrong this often.'"

[15] "In plea bargaining, the minimum is the thing. In this respect, their observations were identical to those of their counterparts in every other city I had visited--save only a few cities in which the minimum and maximum terms of imprisonment are not determined by the trial courts. Given the crowded conditions of most penitentiaries and the customary practices of most parole boards, the attorneys report that the minimum term is usually the critical factor in determining the date of an offender's release. In the vast majority of cases, the maximum term is simply irrelevant. Attorneys in some jurisdictions estimate that eighty percent of the defendants sentenced to the penitentiary are paroled after their first appearance before the parole board. It is, of course, the minimum term that determines the date of this initial appearance. When I repeated this analysis to the prosecutor, he replied that I had simply had made his point for him. Defendants care about when they will hit the streets, and the State's Attorney cares about a sentence that looks severe. The low-minimum, high maximum sentence thus satisfies the interests of both sides. If a paroled offender commits a serious offense after his release, the prosecutor notes, his premature release cannot become the State's Attorney's political problem. The plea agreement that the prosecutor secured gave the parole board authority to hold the defendant for an adequate period and if the board did not choose to exercise that power--if only to free the offender's bed for an incoming prisoner--the political responsibility is theirs." From Alschuler, pp. 108-109. The opposing viewpoint can be found in Newman, pp. 184-85: "Where the realistic sentencing issue is not a choice of incarceration or probation but only a question of the length of time to be served in prison, judges in Michigan routinely reduce charges, pointing out the actual time served will probably be the same on the lesser as on the greater charge. Michigan judges must impose the statutory maximum but have indiscretion as to the minimum. The primary concern of a defendant in Michigan is to avoid a long maximum, and judges are willing to comply since they can compensate

for this apparent sentencing leniency by setting what they consider an appropriate minimum. This type of reduction is common, virtually routine, where the charge concession really does not cost anything. As long as the sentencing judge can still give what he feels is an appropriate sentence and as long as parole eligibility remains or can be fixed at about the same time, the lesser charge brings all of the benefits of the guilty plea without interfering with the sentence. This is an interesting example of the dual perspective of judge and defendant in plea negotiation. The defendant responds to the maximum; he seeks to lessen the most that can happen to him. The prosecutor and the court, however, attuned to the processing of hundreds of cases and perhaps more knowledgeable of current parole practices, negotiate from what might be called practice "norms" of sentencing."

# CHAPTER II

## CONTRIBUTIONS BY ECONOMISTS TO THE THEORY OF COURTS AND PRETRIAL NEGOTIATIONS

The first significant effort to model that portion of the court system involving the prosecutor was undertaken by William Landes.[1] Earlier efforts by economists included investigations of the optimal level of resources devoted to the prevention of crime,[2] the social optimization of the bail system,[3] and participation in illegitimate activities.[4] Of these works, only the effort of Landes has direct bearing upon the analysis of courts presented here. More recently, work by Lachman,[5] Rhodes,[6] Forst and Brosi,[7] and Adelstein[8] has investigated the pattern of trial versus negotiated plea in the court system.

To properly appreciate much of the pioneering and still relevant work in plea bargaining, it is necessary to have a rudimentary understanding of the Expected Utility Hypothesis.[9] This chapter is broken up into two sections. The first section is a statement of the expected utility hypothesis and the formuation of a basic bargaining model given uncertainty. The second part will show how much of the work done on plea bargaining fits within the expected utility/ bargaining framework.

## 2.1 Expected Utility[10]

For a moment, assume that there are two possible states of the world relevant to individual X. If state one occurs he receives a bundle of goods or wealth $W_1$. If state two occurs instead of state one he receives bundle $W_2$. $W_1$ and $W_2$ may consist of money, bundles of goods, or even further possible states of the world with other bundles of goods to be received. Let state one have the probability p of occurring while state two can occur with probability (1-p). Individual X therefore receives $W_1$ with probability p and $W_2$ with probability (1-p).

Before proceeding, it is necessary to make some assumptions concerning the individual's preference for uncertain bundles. In the following let $\succsim$ stand for "at least as preferred as." Let the set $\Omega = \{ W_i / W_i$ is contained in the set of uncertain bundles $\}$.

Assumptions[11]

$W_i \succsim W_j$ is a complete preordering over $\Omega$.[12] (A.1)

$W_i \succsim W_j \Rightarrow W_i \succsim \alpha W_i + (1 - \alpha)W_j$ for all $\alpha \in [0,1]$. (A.2)

$\alpha W_i + (1 - \alpha)W_j \succsim (1 - \alpha)W_j + \alpha W_i$ and (A.3.1)

$(1 - \alpha)W_j + \alpha W_i \succsim \alpha W_i + (1 - \alpha)W_j$

$\alpha(\beta W_i + (1 - \beta)W_j) \succsim \alpha\beta W_i + \alpha(1 - \beta)W_j$ and (A.3.2)

$\alpha\beta W_i + \alpha(1 - \beta)W_j \succsim \alpha(\beta W_i + (1 - \beta)W_j)$

Assumption A.2 states that if $W_i$ is at least as preferred as $W_j$ then any chance at obtaining $W_i$ is preferred to $W_j$ with certainty.

Assumption A.3.1 states that the order of combinations is irrelevant while A.3.2 states that the net probabilities of receiving an outcome are important in its valuation and that this valuation is not affected by the sequencing of separate probabilities.

Given these three assumptions it is possible to define a utility function on $\Omega$ such that $U(pW_i + (1 - p)W_j) = pU(W_i) + (1 - p)U(W_j)$.[13] This is the so-called Expected Utility Hypothesis.

## 2.2 The Individual in the World of Uncertainty

Suppose an individual derives utility from wealth W and that utility increases with increases in W. Assume that the individual is faced with the following choices: (1) to receive $W_2$ with certainty or (2) to receive $W_1$ with probability $p_1$ or $W_3$ with probability $(1 - p_1)$ where $W_1 < W_2 < W_3$. Using the Expected Utility Hypothesis the certainty option is preferred if $U(W_2) > p_1U(W_1) + (1 - p_1)U(W_3)$. Choice two is preferred if the inequality is reversed and the individual is indifferent between the choices if the inequality is replaced by an equality. The expected value of choice two is $p_1W_1 + (1 - p_1)W_3$. An individual is said to be risk averse if he always prefers to obtain the expected payoff of the risky choice to the risky choice itself; i.e., if

$$U(p_1W_1 + (1 - p_1)W_3) > p_1U(W_1) + (1 - p_1)U(W_3) \tag{1}$$

then the individual is risk averse. It can be shown that if U is concave that the inequality in equation 1 will hold for any $p \in [0,1]$.

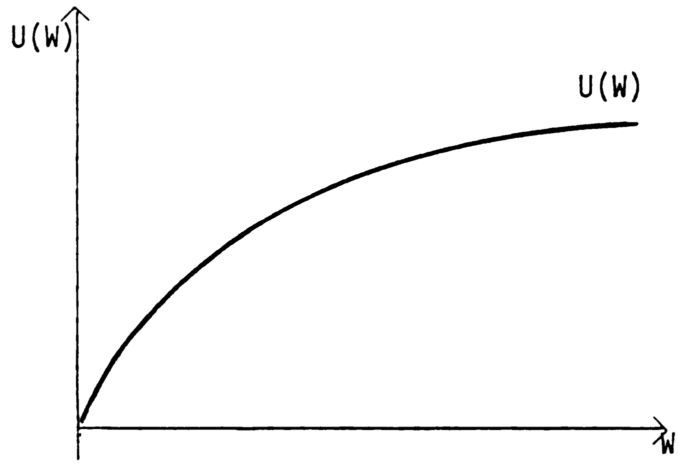



Figure 1.--Example of utility function of risk averter.

It follows from equation 1 that a risk averse individual is willing to pay something in order to avoid risky situations. Thus when equation 1 holds it is possible to find values of $W = W_k$, such that $W_k < p_1W_1 + (1 - p_1)W_3$ and $U(W_k) > p_1U(W_1) + (1 - p_1)U(W_3)$. Let $W_o$ be that W such that $U(W_o) = (1 - p_1)U(W_3) + pU(W_1)$. The individual is then willing to pay up to $W_L = p_1W_1 + (1 - p_1)W_3 - W_o$ to avoid the risky choice in the sense that for any $W_z \geq W_L$ it holds that $U(W_z) > p_1U(W_1) + (1 - p_1)U(W_3)$.

Given that when faced with uncertainty an individual is willing to give up something in order to eliminate that uncertainty opens up the possibility, among other things, of bargaining. Thus for example two individuals (assume prosecutor and defendant) when engaged in a conflict (trial) with an uncertain outcome both can gain by the elimination of uncertainty.

Assume that there are two states of the world $S_1$ and $S_2$ with respective probabilities p and (1 - p) and associated levels of wealth

$W_1$ and $W_2$. Given p the individual's expected utility will be a function of the wealth he will have in each of the two states. Thus $U(W_1,W_2) = pU(W_1) + (1 - p)U(W_2)$. This relationship may be related to standard indifference curve analysis. The indifference curve II' for example in Figure 2 below represents all those combinations of $W_1$ and $W_2$ where total expected utility does not change.

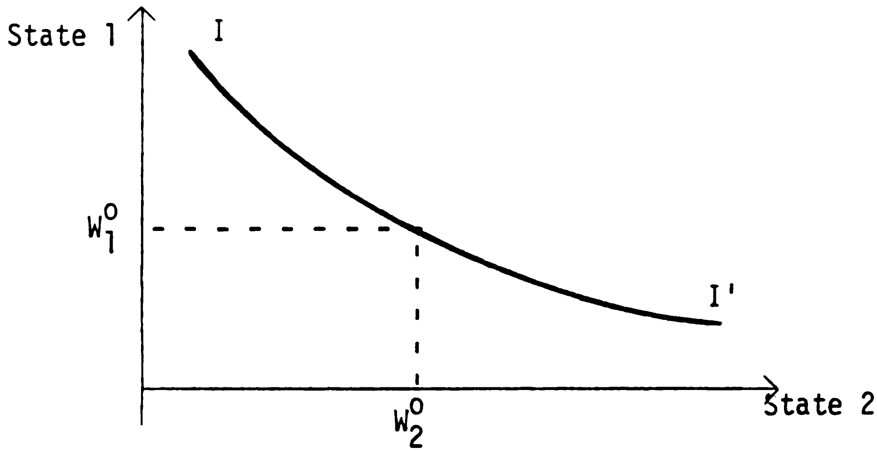



Figure 2.--Combinations of $W_1$ and $W_2$ where total expected utility is a constant.

The indifference curve in Figure 2 was drawn upon the assumption that $U'(W) = dU/dW > 0$ and $U" = d^2U/dW^2 < 0$.[14] The slope of the indifference curve II' at $W_1^o$, $W_2^0$ is

$$dW_1/dW_2 = \frac{-(1 - p)U'(W_2^o)}{pU'(W_1^o)} \quad .$$

This equation states how much $W_1$ must be given up if $W_2$ is to increase by a small amount and the individual is to remain at the same level of total expected utility. Two initial observations can be made about $dW_1/dW_2$. First if $W_1^o > W_2^o$ the price of $W_2$ in terms of $W_1$ is higher than if $W_2^0 > W_1^0$. This follows directly from the assumption of decreasing marginal utility. Secondly since $d(1 - p)/dp < 0$ the price of $W_2$ in terms of $W_1$ increases as p falls. Thus as state 1 becomes less certain the individual is willing to give up more of $W_1$ for a given increase in the now more certain $W_2$.

## 2.3 Bargaining

In Figure 3 below the possibility of trading between two individuals is analyzed. Starting at the point $O_1$ and looking in a northeast direction we see indifference curves for individual one. Starting at point $O_2$ and looking in a southwest direction we see the indifference curves for individual two.

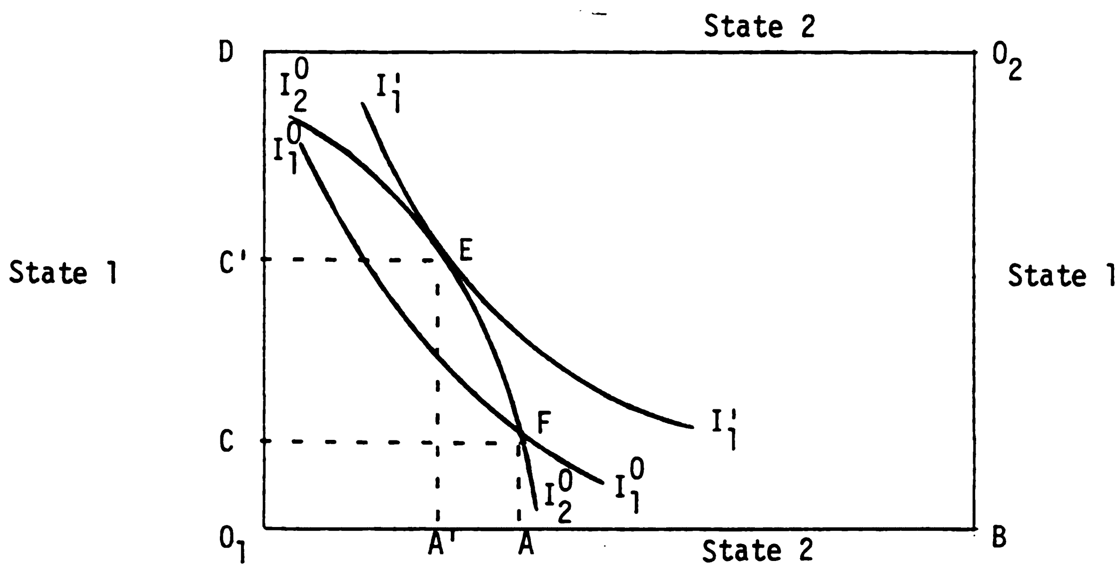



Figure 3.--Bargaining space.

Let the point labeled F represent the initial allocation of resources between the two individuals. Individual one has wealth A in state 2 and C in state 1 while individual 2 has W-A wealth in state 2 and W-C wealth in state 1, where W is the total level of wealth to be divided. It is easy to show that the initial allocation of resources is not Pareto optimal.[15] For example assume that the wealth is redistributed through bargaining to that represented by the point E. At this point individual one is on a higher indifference curve than previously and thus better off while individual 2 has not suffered by the trade since he is still at the old level of utility. The only point where the two individuals cannot profit from trade is at a point like E where the indifference curves of the two individuals are tangent. At this point of tangency (and optimality) it is true that

$$\frac{-(1 - p_1)U_a'(A')}{p_1 U_a'(C')} = \frac{-(1 - p_2)U_b'(W-A')}{p_2 U_b'(W-C')} , \qquad (2)$$

where $p_1$ is the first individual's assessment of the likelihood that state 1 will occur. Likewise $p_2$ is the estimate of individual 2 that state one will occur. $U_a$ is the utility function of individual 1 and $U_b$ is the utility function of individual 2.

Let the wealth of individual one in state one be $W_1^1$ and in state two be $W_2^1$, where $W_1^1 < W_2^1$. Let individual one's assessment of the likelihood of state one occurring be $p_1$. Let the wealth of individual two in state one by $W_1^2$ and in state two be $W_2^2$ and his assessment of state one occurring $p_2$, where $W_2^2 > W_1^2$ and $W_1^1 < W_2^1$ and

$W_1^1 + W_1^2 = W_2^1 + W_2^2 = W$. The conditions on W imply that what individual one loses two wins and vice versa.

Individual one's expected utility function is therefore $p_1U^1(W_1^1) + (1 - p_1)U^1(W_2^1)$ and that of individual two is $p_2U^2(W_1^2) + (1 - p_2)U^2(W_2^2)$. The necessary condition for Pareto optimality is

$$\frac{-(1 - p_1)U^{1'}(W_1^1)}{p_1U^{1'}(W_2^1)} = \frac{-(1 - p_2)U^{2'}(W_1^2)}{p_2U^{2'}(W_2^2)} \quad . \tag{3}$$

If both individuals are risk averse the optimality condition in 3 can occur only if $p_1 > p_2$. That is, individual one's assessment of winning must be greater than individual two's assessment of individual one winning. To show this assume that $p_1 = p_2$ and notice that

$$\frac{U^{1'}(W_1^1)}{U^{1'}(W_2^1)} > 1 \quad \text{since } W_1^1 < W_2^1 \text{ ,}$$

and

$$\frac{U^{2'}(W_1^2)}{U_p^{2'}(W_2^2)} < 1 \quad \text{since } W_1^2 > W_2^2.$$

The only way equality can be obtained is when $p_1 > p_2$ since as $p_1$ increases $(1 - p_1)/p_1$ decreases. Note that this means that if the two individuals have the same assessment of the probabilities the only way to obtain Pareto optimality is through side payments. In the case of criminal proceedings through the prosecutor's office, this would occur through negotiations. In this case where $p_1 = p_2$ a transfer of wealth would occur (if the Pareto optimum is to be reached) where

$W_1^1 = W_2^1$ and $W_1^2 = W_2^2$. That is, each individual would hedge so that his winnings in both states were equal.

## 2.4 Negotiated Plea Theory and the Expected Utility Framework

The initial pioneering work on negotiated pleas is by Landes.[16] In this work the defendant is seen as having two mutually exclusive states. In state one the defendant goes to trial and is convicted. In state two he goes to trial and wins. In the first state his wealth is $Wc = W - sS - rR$ and in the second state his wealth is $Wn = W - rR$ where

W is initial wealth,

S is length of sentence received by the defendant if convicted,

s is the present value of the average cost of losses per unit of S,

R is the input of resources to the case by the defendant,

r is the average price of a unit of R.

If $p_d$ is the defendant's estimated probability of winning the case, then the expected utility function that he wishes to maximize is

$EU = (1 - p_d)U(W - sS - rR) + p_dU(W - rR) = (1 - p_d)U(Wc) + p_dU(Wn).$

The rate at which a unit of Wc can be obtained by giving up some Wn and still maintain the same level of expected utility is

$$dWn/dWc = \frac{-1(1 - p_d)U'(Wn)}{p_dU'(Wc)} \quad \text{where } \frac{U'(Wn)}{U'(Wc)} < 1 \text{ since } Wn > Wc.$$

As the probability of acquittal falls, the absolute value of dWn/dWc gets larger. Thus the individual is willing to give up more Wn for a one unit increase in Wc (wealth given conviction) as the probability of conviction rises.

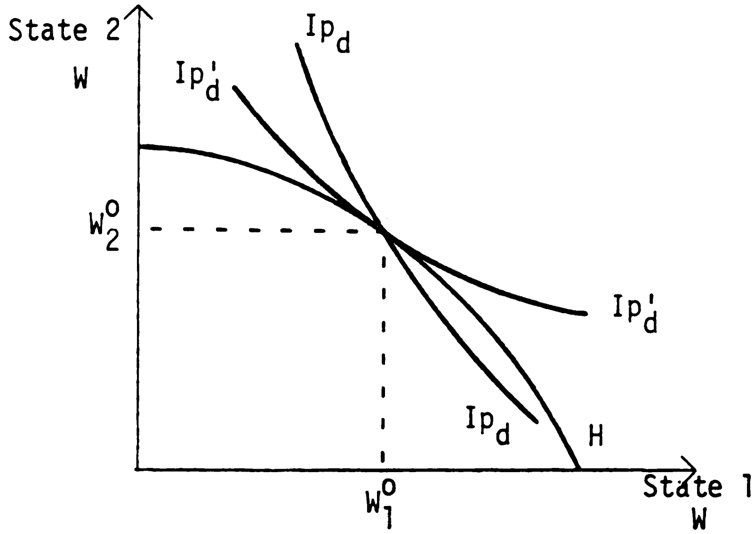



Figure 4.--$p_d < p'_d$.

Figure 4 depicts the actions of the defendant given an increase in the probability of conviction. Initial optimal allocation for the defendant occurred at the point $(W_1^o, W_2^0)$ where his indifference curve $I_{p'_d}$ is just tangent to his resource trade-off curve H. An increase in the probability of conviction (a decrease in $p_d$) results in a new set of indifference curves. For the new set the derivative dWn/dWc at the point $(W_1^o, W_2^o)$ is more negative than the slope of H. Since indifference curves from the family cannot cross, this implies that the new equilibrium will occur at a point $(W_1^1, W_2^1)$ where $W_1^1 > W_1^o$ and $W_2^1 < W_2^o$. In other words, the defendant will spend less on "beating the rap."

The conceptualization of the prosecutor is of a simpler form. He is seen as maximizing $p_xS_x + p_oS_o$ subject to a linear budget constraint. In this formulation $p_x$ is the probability that the defendant of interest above is convicted. $S_x$ is the sentence the defendant will receive upon conviction. $p_o$ and $S_o$ are composite variables defined like $p_x$ and $S_x$ but for all other defendants. It is assumed that $p_x$ is a function of resources spent, evidence, etc.

An explicit theory of how the prosecutor develops the level of concessions that he is willing to make is not developed except in the most general intuitive fashion. Bargaining increases $p_x$ to 1 and allows a reallocation of resources to other cases. This would lead to an increase in total utility if $s_x$ were to remain the same. The maximum level of concessions $\Delta s$ that the prosecutor is willing to make with the defendant and still retain his initial level of total utility defines a new reduced charge against defendant X labeled $S_p = S_x - \Delta S$. $\Delta S$ will increase as the savings from negotiation increase. Finally, Landes notes that if bargaining is to occur both parties must benefit from it. If $S_p$ is the minimum sentence that the prosecutor offers the defendant and $R < R$ is the transaction cost of accepting this bargain then a sufficient condition for bargaining to occur is $sS_p < r(R - \hat{R})$ since this guarantees that $W - sS_p - r\hat{R} > W - sS_x$.

Landes uses cross-section data from state and federal courts, regression analysis, and a reduced form model that includes independent variables adduced from the theoretical model as significant in determining whether a case is decided by trial or negotiated plea to derive the following results:

1. a. The more risk averse[17] the defendant, the more likely a settlement will be reached.

b. The greater the probability of conviction at trial, as viewed by the defendant, the more likely the defendant is to negotiate.

2. Denial of bail increases the probability of settlement via negotiated plea rather than trial.

3. As the cost differential between plea bargaining and trial rises, the propensity to settle without trial increases.

The theory of Landes can be put into the general framework of section one and the same type of conclusions will result. Assume that the prosecutor has a p probability of getting defendant X convicted of a charge with sentence $Z_1$. The case against defendant X can be lost with probability $(1 - p)$. The sentence defendant X then receives is $Z_2$. $Z_2$ is assumed to be zero or near zero. A nonzero value for $Z_2$ should be interpreted as the time defendant X spends in jail prior to trial. The prosecutor is assumed to maximize his expected utility which is a function of p, $Z_1$ and $Z_2$. That is, the prosecutor seeks to maximize $U = pU(Z_1) + (1 - p)U(Z_2)$ subject to a budget constraint which is $H(Z_1,Z_2,R) = 0$ where R is resources.

The defendant perceives that he can be convicted with probability k. In this case his total freedom is $T - Z_1$, where $T > Z_1$ is the length of the defendant's time horizon. The defendant believes that he has a $(1 - k)$ chance of acquittal and in this case his freedom is $T - Z_2$. The defendant is assumed to maximize $U = kV(T - Z_1) + (1 - k)V(T - Z_2)$ subject to an appropriate budget constraint (which is $G(Z_1,Z_2,W) = 0$ where W is wealth. As shown in section 2.3 of this chapter, if the prosecutor and defendant are to reach a settlement via trial then it is necessary that

$$\frac{pU'(Z_1)}{(1 - p)U'(Z_2)} = \frac{kV'(T - Z_1)}{(1 - k)V'(T - Z_2)} \tag{4}$$

Landes states, "If both parties agree on the probability of conviction by trial, a settlement will take place."[18] This follows from equation 4. Setting p = k leads to

$$\frac{U'(Z_1)}{U'(Z_2)} < \frac{V'(T - Z_1)}{V'(T - Z_2)}$$

since $U'(Z_1) < U'(Z_2)$ and $V'(T - Z_1) > V'(T - Z_2)$ since $(T - Z_1) < (T - Z_2)$ or $Z_1 > Z_2$.

Also Landes states, "Suppose the prosecutor and defendant differ in their estimates of the trial conviction probability, if $p < k$ a trial becomes an even less favorable gamble in comparison to $p = k$."[19] This follows from equation 4 since for a given k the right hand side decreases in value as p falls from $p = k$. Also note that equation 4 implies that if the case is to be settled in court via trial then a necessary condition is that $p > k$. In other words, the prosecutor's estimate of conviction must be greater than that of the defendant. Finally, note that an increase in $Z_1$, the sentence received upon conviction, decreases the probability of settlement via trial since the inequality in equation 4 becomes more pronounced since $T - Z_1$ decreases in value. Thus the more the defendant has to lose, <u>ceteris paribus</u>, upon conviction the less likely the case is to be resolved via trial. Or viewed slightly differently, the higher $Z_1$ the lower the defendant's estimate of conviction must be if the case is to be resolved via trial. It does not appear that Landes made this last observation from his theory and he seems not to have understood this phenomenon since later in his path-breaking paper he worries over the apparent anomaly this presents: The regression coefficients on the sentence variable S (sic $Z_1$) do not support the hypothesis that the likelihood of trial is greater for defendants accused of crimes carrying longer sentences. The above confusion stems from Landes' missed guess on the direction of an inequality and his dismissing of the data rather than his hypothesis. The theory within the main body of this dissertation develops a rationalization for why an increase in the severity of sentence, ceteris paribus, need not lead to an increase in the demand for trial and in fact can lead to the pressing of reduced charges by the prosecutor.

Landes' work centered around detailed analysis of the individual prosecutor and defendant. William Rhodes extended Landes' model so that determination of the number of cases handled via plea negotiation vis-à-vis trial could be determined.[20] Rhodes' intent was to develop and extend Landes' model to clarify the relationship between individual decision makers and the overall mechanics of the court.

Rhodes follows Landes and assumes that the prosecutor's probability of convicting defendant $X_i$ at trial is a function of public expenditures on the case, $R_i$, defendant's expenditures on defense, $R_i^*$, and other measures of the defendant's convictability, $Z_i$. Thus $P_i = P_i(R_i, R_i^*, Z_i)$. It is assumed that there are only two levels of expenditures possible on case $X_i$; $c_t$ is spent if the prosecutor pursues the trial option while $c_s$ is spent if the negotiation route is selected. In other words, $P_i$ can be viewed as exogenous since the levels of $c_t$ and $c_s$ are not determined from within the model. The probability thus of convicting defendant $X_i$ via trial is $P_i = P(c_t, R_i^*, Z_i) = P(X_i)$. It is further assumed that when sufficient funds are not available for trial prosecution of all defendants that they are prosecuted in the order of convictability. The defendants are assumed ranked in order from 1 to N and that $\partial P/\partial X < 0$; i.e., as the prosecutor moves down the list of defendants the probability of conviction falls. It is assumed, based upon the theoretical work of Landes, that as the cost of trial to the defendant increases, the number of defendants seeking resolution of their case by way of trial will fall. It is also assumed that as the level of concessions for plea bargaining is increased, the number of defendants seeking negotiated pleas will increase. In functional form, the fraction of defendants t going to trial can be written as:

$$t = t(\Delta, B) \text{ where}$$

$\Delta$ = measure of concessions given by the prosecutor in return for the defendant entering a guilty plea,

B = defendant's overall cost of going to trial. From the above assumptions about the effect of concessions and costs of trial on the number of defendants seeking trial, $\partial t/\partial \Delta < 0$, $\partial t/\partial B < 0$. Further, it is assumed that $\partial^2 t/\partial B^2 > 0$.

The prosecutor's gain from trial is[21]

$$t(\Delta, B)\int_0^n P(X) S dX ,$$

i.e., the prosecutor's gain from trial is the expected total sentence of convicting all defendants by trial times the fraction of defendants who are brought to trial. S is the average sentence to be received by all defendants. The gain to the prosecutor from negotiation is $(1 - t(\Delta,B))\int_0^n (P(X)S - \Delta)dx$, i.e., the prosecutor's expected gain from guilty pleas equals the sum of the expected plea bargaining sentences for all defendants times the fraction of defendants who plead guilty. The sum of the two partial gains gives the prosecutor's total gains,

$$t(\Delta,B)\int_0^n P(X)Sdx + \int_0^n P(X)Sdx - n\Delta - t(\Delta,B)\int_0^n P(X)Sdx - nt\Delta = \int_0^n (P(X)S - \Delta)dx - n\Delta t(\Delta,B). \tag{5}$$

The prosecutor is assumed to behave as if he maximized equation 5 subject to the budget constraint,

$$R = t(\Delta,B)nc_t + (1 - t(\Delta,B))nc_s, \tag{6}$$

where R is the total budget of the prosecutor.

Maximization of equation 5 subject to the budget constraint of equation 6 results in the first order conditions necessary for an interior maximum of

$$PS - (1 - t)\Delta - \lambda(t(c_t - c_s) + c_s) = 0$$

$$-(1 - t)n + (\partial t/\partial\Delta)n\Delta - \lambda(\partial t/\partial\Delta(c_t - c_s)n) = 0,$$

$$R - tn(c_t - c_s) - nc_s = 0,$$

where $\lambda$ is the Lagrangian multiplier.

Rhodes shows that $\partial n/\partial R > 0$,

$$\partial \Delta/\partial R < 0,$$

$$\partial n/\partial B > 0.$$

The first inequality implies that as the resources available to the prosecutor are increased the number of prosecutions will increase. The second inequality implies that with an increase in resources the prosecutor will offer fewer concessions in return for a guilty plea. The last inequality states that as the cost of going to trial increases for defendants the prosecutor will be able to handle more cases since fewer defendants will be seeking resolution of their case via trial. Thus the average cost per defendant falls, making it possible for the prosecutor to handle more cases.

Rhodes backs up his theoretical work with statistical cross-section analysis of the federal courts. At least two criticisms can be leveled against the work of Rhodes. First, the theory implies that defendants do not react to the actions of the prosecutor since $R_i^*$ is assumed fixed throughout the analysis. Secondly, if $c_t$ is not held constant the determinancy in equations a - c no longer holds. For example, if $c_t = \theta(t)$, i.e., the average amount of resources spent per trial is a function of the number of defendants who will be tried, it is not hard to show that most of the results of Rhodes do not follow unambiguously. Rhodes is misleading when he remarks: "Strictly speaking the prosecutor's expenditures on trials should be treated as endogenous to the model. However, it is difficult to solve the general equilibrium solution analytically with an endogenous expenditure specification."[22] It is unfortunate that Rhodes leaves the reader with the impression that with $c_t$ endogenous his results still necessarily follow.

Rhodes in a later work[23] develops some empirical estimates of the effects of initial plea, final plea, bail involving money, bail not involving money, age, sex, race, multiple charges, and judge on the probability of receiving a prison sentence upon conviction. The findings of this study (based upon the same data used in the prior study but with different econometric specification) support the earlier findings. The most important being that the percentage of defendants pleading guilty increases as the level of concessions granted for pleading guilty is increased, while the major determinant of concessions is the size of the prosecutor's caseload.

Forst and Brosi attempt to extend Landes' single period analysis to a multiperiod framework.[24] They develop their model as follows. Assume that in the $t^{th}$ time period, $n_t$ cases are brought to the prosecutor for disposition. Let $P_{it}$ be the probability of convicting via trial the $i^{th}$ defendant in time period t. Following Landes it is assumed that $\partial P_{it}/\partial R_{it} > 0$ where $R_{it}$ is the amount of resources expended by the prosecutor on case i in the $t^{th}$ time period. If $S_t$ is a measure of the seriousness of crimes committed in time period t then the present value of crimes committed through time period $t + k$, $k > 0$, is

$$\Sigma \frac{S_t}{(1 + d)^t}$$

where d is the prosecutor's discount rate. It is assumed that the prosecutor is indifferent between all streams of crimes with the same present value. For the $i^{th}$ defendant the prosecutor can envision a stream of crimes which could be committed by the defendant if he is not convicted of the current charge,

$$S_{i1}, S_{i2}, \ . \ . \ . \ , S_{it} \tag{7}$$

If the defendant is convicted of the current charge the prosecutor envisions another different stream of crimes

$$S^c_{i1}, S^c_{i2}, \ . \ . \ . \ , S^c_{it} \tag{8}$$

where the difference of the present value of the two streams is

$$D_i = \sum_{k=1}^{t} \frac{(S_{ki} - S^c_{ki})}{(1 + d)^k} , \tag{9}$$

which represents the present value of crimes that are averted by the conviction of defendant i. It is assumed that $D_i$ depends upon the characteristics of the individual $X_i$ and the severity of punishment $T_i$ for the current crime; thus $D_i = D_i(H_i,T_i)$ where $\partial D_i/\partial H_i > 0$ and $\partial D_i/\partial T_i > 0$. The only relevant variable in $H_i$ is taken to be the defendant's criminal history.[25] $\partial D_i/\partial H_i > 0$ implies that as the measure of criminal activity increases, a larger increase in the present value of averted crimes will occur with incarceration of the defendant since a defendant with a large $H_i$ is more likely to commit a crime if set free. Two models of the prosecutor's utility function are presented. The first:

$$U = U(\Sigma D_i) + \lambda(B - \Sigma R_i) \tag{10}$$

which is explicitly stated as:

$$\max \quad \Sigma P_i D_i + \lambda(B - \Sigma R_i) \ . \tag{11}$$

Equation 11 is consistent with equation 10 but does not necessarily follow from it as the authors seem to imply.[26] The authors reject this model formulation a priori for being too limited in that it excludes $T_i$, the punishment variable. Thus the prosecutor is seen not only as seeking to minimize the present value of future crime but also to maximize the punishment for that crime. This brings about the second model:

$$\max U = U(\Sigma T_i, \Sigma D_i) + \lambda(B - \Sigma R_i) \tag{12}$$

The authors say, "Under this specification, the prosecutor will maximize expected utility, subject to his budget constraint, as given by $E(U) = \Sigma P_i U(T_i, D_i) + \lambda(B_i - \Sigma R_i)$."[27] This, if based upon equation 12, implies a violation of the linear transformation rule and includes a myriad of zero restrictions on variable coefficients among other things.

The authors use two stage least squares to estimate

$$R = a + b_1 P + b_2 T + b_3 H, \text{ where}$$

R = number of days the prosecutor carries the case,

P = probability in the case that the defendant will be convicted,

T = Sellin-Wolfgang index of crime seriousness, and

H = number of known prior arrests.

In the first stage of the estimation process, P is estimated by regressing it on a full set of exogenous variables which include evidence, age, type of crime, number of co-defendants, etc. They do not state so, but it is implied that P is simply dichotomous, zero or one, depending upon whether the defendant is set free or convicted.

The theory developed implies that the appropriate P is that of the prosecutor. The P is estimated via a reduced form structure. The authors have no discussion of whether the P that is estimated reflects the beliefs of the prosecutor or those of the defendant. The list of exogenous variables used to estimate P would seem to imply that the defendant's estimate of the probability of conviction is being estimated since the exogenous variables would appear to be those that would shift the prosecutor's estimate of the probability of conviction and thus trace out the defendant's probability of conviction curve. This may not be a serious error, since the prosecutor's and defendant's estimates of the probability of conviction are most likely highly correlated, but it might partially account for the result discussed below. The authors have one unique result, namely, that prosecutors are more sensitive to changes in evidence than to changes in crime severity. The prosecutor is seen as pursuing the cases he can win and not necessarily defendants who have commited serious crimes.

All of the above authors have assumed that the prosecutor maximizes the sum of expected sentences and are thus partially outside a general theory of expected utility maximization. Only Landes develops a theory of both the prosecutor and defendant.

Richard Adelstein has addressed himself partially to these objections.[28] In his theory he assumes that the prosecutor in each case faces an expected utility function of the form $U = PU(S)$ where S is the expected sentence the defendant will receive if convicted and P is the probability of conviction. P is assumed a function of time (t) and resources spent (R) on the case where

$$(a)\ \partial P/\partial R > 0,\ \partial^2 P/\partial R^2 < 0 \ \text{ and}$$

$$(b)\ \partial P/\partial t < 0 \ \text{ and } \partial^2 P/\partial t^2 > 0.$$

The conditions in (a) assert that as the prosecutor increases the level of resources devoted to the case the probability of conviction increases but at a decreasing rate. The conditions in (b) imply that the passage of time hurts the prosecutor's chances of getting a conviction and also that the case deteriorates more quickly in the early stages of prosecution because of the passage of time than it does in later stages of adjudication. Other authors have given notice to the influence of time but have not explicitly incorporated it into P.[29] Landes, among others, however, has assumed that P is a function in part not only of resources devoted to the case by the prosecutor but also by the defendant.[30]

The prosecutor's problem is to allocate the agency's budget so as to maximize the sum of the expected utilities over the anticipated caseload. This highlights the first weakness of this theory, as with the rest of the previous works, that total utility is simply the sum of the utilities over the individual cases. This assumption is not an application of the expected utility hypothesis discussed previously.

Adelstein gives what can be considered the most explicit statement of prosecutorial costs in any published work.[31] He breaks the costs of pursuing a case into two types. First, there are the outlays in preparing the case for trial. These outlays are used in an effort to increase P, the probability of conviction. The second category of costs are expenditures on pretrial detention of defendants not released on bail. He notes that typically these costs are not directly subtracted from the prosecutor's budget but are subtracted from the budget of the corrections department. This department competes with the prosecutor for operating revenue from a limited governmental allocation for preconviction criminal justice services. In either case, the more spent on pretrial detention the less that remains in the total governmental budget for trial preparation. The detention cost is formalized as $D = \int_0^t D(Z)dZ$ where D(.) is a function yielding detention costs per day. If $B_j$ is the total expenditure on the $j^{th}$ case and $\Sigma B_i = B$ is the total prosecutorial budget, then the optimization problem of the prosecutor consists in finding $\Sigma B_i^* = B$ such that total expected utility is maximized. Each $B_i^*$, $i = 1, \ldots, m$ represents the optimal allocations over the caseload given that each case is resolved by way of trial. If the prosecutor initiates plea bargaining in case 1, then there is a savings of $w_1$. The prosecutor then maximizes $V_1(P,w_1)$ subject to the constraint $B_1 + w_1 \leq B_i^*$. $V_1$ is a utility function of the form $V(PU(S), w_1)$. If $w_1 > 0$ in the optimal solution, then the prosecutor will seek resolution of the case via bargaining. Only when $w_1 = 0$ will the prosecutor seek settlement by way of trial. This optimization process is done sequentially for each case until a maximum is reached. Adelstein does not anticipate that the prosecutor reaches the maximum since he most likely goes through the process described only once. However, he does believe that it captures the essence of the process used by the prosecutor.[32] A criticism can be raised concerning the present formulation. Adelstein argues that in the second round the prosecutor maximizes $V(PU(S),w_1)$ subject to $B_1 + w_1 < B_1^*$. However, given plea bargaining, $P = 1$. Thus in searching for a plea bargaining solution the search is not over w and thus $P = P(w)$ but rather over S. The prosecutor seeks that $\bar{S}$ such that $U(\bar{S}) = PU(S^*)$. $S^*$ is the optimal S in the initial problem. This $\bar{S}$ is the smallest S that the prosecutor is willing to offer during plea negotiations. Following the procedure set down by Adelstein, the prosecutor would then maximize $\Sigma P_i U(S_i)$ subject to $\Sigma B_i^* = B - \bar{w}_1$ where $\bar{w}_1$ is the cost associated with plea bargaining in case 1. This then produces a new set of budget allocations for resolution via trial in the m - 1 remaining cases. The prosecutor then determines $\bar{S}_2$ if it exists and then the process starts all over and continues until an expected utility maximum is reached. It is not necessary that such a simple search technique produce the maximum. Adelstein further assumes that the defendant wishes to minimize a loss function

$$\min L = \underline{P}L(S) \text{ subject to } B(t) = X - C_b + \int_0^t E(Z)dZ$$

where $\underline{P}$ is the defendant's estimate of the probability of conviction, X is the initial wealth of the defendant, $c_b$ is the cost of bail, $c_b = 0$ if bail is denied, and $\int_0^t E(Z)dZ$ is the earnings of the defendant after arrest but prior to trial. $\int_0^t E(Z)dZ$ is assumed zero if bail is denied. As in the case of the prosecutor $\underline{P} = \underline{P}(B,t)$, $\partial P/\partial B < 0$, $\partial^2\underline{P}/\partial B^2 > 0$ and $\partial\underline{P}/\partial t < 0$ and $\partial^2\underline{P}/\partial t^2 > 0$. A boundary solution with $t = T^*$, $T^*$ the trial date implies that the defendant seeks resolution of the case via trial. For any $t < T^*$ the defendant maximizes utility by negotiating with the prosecutor. Denote the solution to this optimization problem as $(P^*,t^*)$. Let $S_{min}$ be that punishment associated with full investment of the defendant's resources in sentence reduction via trial. Then for every $t^* < T$, the convexity of the loss function ensures the existence of a subset of feasible points $(S,t)$ such that $L(S^*,t) < L(S,t) < L(S_{min},0)$. This subset of punishment/pretrial time pairs defines the limits of the defendant's bargaining set. It would seem that this is not an adequate formulation unless $S_{min} = 0$ since the defendant is always willing to settle for no punishment.

In the development of the bargaining theory the above theory has two functions. First, it acts as an existence theorem and rationale for bargaining. Second, it establishes the importance of time and time-dependent optimality solutions.

In specifying the bargaining process, Adelstein develops a variant of the Cross bargaining model.[33] Let $P_p(t)$ be the utility-maximizing punishment of the prosecutor for defendant X. Then the value of a given plea agreement to the prosecutor may be written as

$$U(P_p,v) = U(P_p) - \int_0^v D(t)dt$$

where v is the prosecutor's estimate of the time to settlement. D(t) should be expressed in utility units if consistency is to be maintained, but the author repeatedly attributes a dollar cost to D(t). Specifying D(t) as representing a utility measure of cost rather than a monetary measure does not change the specification of the model at this level of abstraction. Let

$$v = \frac{P_p(t) - P_d(t)}{r_d(t)}$$

where $P_d(t)$ is the defendant's utility-maximizing level of punishment and $r_d(t)$ is the rate of concession of the defendant as perceived by the prosecutor. If $D(t) = Dv$, D a constant, then the prosecutor's optimal initial offer is the solution to the following equation, which is maximized with respect to $P_p(t)$.

$$U(P_p,0) = U(P_p) - D\left\{\frac{P_p(0) - P_d(0)}{r_d(0)}\right\} \tag{13}$$

Maximization of equation 1 results in $\partial U/\partial P_p = D'/r_d$. The model of the defendant is symmetric and will not be discussed here. Rules are specified for adjustment to the $r_d$ and $r_p$ functions. A bargain occurs at time $t^*$ when $P_p(t^*) = P_d(t)$ if $t^*$ $T^*$ the date of trial. Adelstein notes that direct estimation would be difficult since utility functions and learning parameters would have to be estimated.[34] No testing of the model or of secondary hypothesis is performed. This model offers little that is new on a conceptual basis to the theory of plea bargaining. Nor does the author suggest a way for overcoming the direct estimation problems of his model.[35]

This dissertation benefits greatly from the previous unpublished work of J. Lachman.[36] Analysis of her work aided in the development of a model specification which was tractable to empirical investigation. In her work Lachman developed a Pareto bargaining set in which a bargained plea must be contained. The device used to construct this set was labeled the switch function by her.[37] The use of this device and the perspective gained by her analysis of shifting opportunity sets aided in the development of this dissertation more directly than the work of any other author. Rather than present a superficial analysis of her work here, reference to her study is made in Chapters III and IV.

## Footnotes--Chapter II

[1]William M. Landes, "An Economic Analysis of the Courts," The Journal of Law and Economics 14 (April 1971): 61-108.

[2]Gary S. Becker, "Crime and Punishment: An Economic Approach," Journal of Political Economy 76,2 (1968): 169-217.

[3]William M. Landes, "Rules for an Optimal Bail System: An Economic Approach," The Journal of Legal Studies 2,1 (1973): 79-106.

[4]Isaac Ehrlich, "Participation in Illegitimate Activities: An Economic Analysis," in Essays in the Economics of Crime and Punishment, ed. Gary Becker and William Landes (New York: Columbia University Press), pp. 68-134.

[5]Judith A. Lachman, "An Economic Model of Plea Bargaining in the Criminal Court System" (Ph.D. dissertation, Michigan State University, 1975).

[6]William M. Rhodes, "The Economics of Criminal Courts: A Theoretical and Empirical Investigation," The Journal of Legal Studies 5 (June 1976): 311-40.

[7]Brian Forst and Kathleen Brosi, "A Theoretical and Empirical Analysis of the Prosecutor," The Journal of Legal Studies 6 (January 1977): 177-92.

[8]Richard P. Adelstein, "The Plea Bargain in Theory: A Behavioral Model of the Negotiated Guilty Plea," Southern Economic Journal 6 (January 1978): 488-503.

[9]John V. Neuman and Oskar Morgenstern, Theory of Games and Economic Behavior (New York: John Wiley and Sons, 1944).

[10]Much of this section is an adaptation of John P. Gould, "The Economics of Legal Conflicts," The Journal of Legal Studies 2 (June 1973): 279-300.

[11]The following description of assumptions can be found in Josef Hadar, Mathematical Theory of Economic Behavior (Reading, Pa.: Addison-Wesley Publishing Co., 1971) and Neuman and Morgenstern, Theory of Games and Economic Behavior.

[12]A description of orderings can be found in Appendix A, the latter part of section A.3.

[13]Hal R. Varien, Microeconomic Analysis (New York: W. W. Norton Co., 1978), p. 106; Neuman and Morgenstern, p. 26.

[14]Hadar, pp. 177-80.

[15]Varien, p. 145.

[16]Ibid.

[17]J. W. Pratt, "Risk Aversion in the Small and the Large," Econometrica 33 (January 1964): 122-36.

[18]Landes, "An Economic Analysis of the Courts," reprinted in Essays in the Economics of Crime and Punishment, ed. Gary Becker and William Landes (New York: Columbia University Press, 1974), pp. 164-214.

[19]Ibid., p. 172.

[20]Rhodes.

[21]All integrals are of the Lesbegue form.

[22]Rhodes, p. 313.

[23]William Rhodes, "A Study of Sentencing in the Hennepin County and Ramsey County District Courts," The Journal of Legal Studies 6 (June 1977): 333-52.

[24]Forst and Brosi.

[25]Adelstein, p. 500.

[26]Forst and Brosi, pp. 182-83.

[27]Ibid., p. 183.

[28]Adelstein.

[29]Lachman, p. 35; Landes, p. 191.

[30]Landes, p. 166.

[31]Adelstein, pp. 490-91.

[32]Ibid., p. 492.

[33]John Cross, The Economics of Bargaining (New York: Basic Books, Inc., 1969).

[34]Ralph Keeney and Howard Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs (New York: John Wiley and Sons, 1976).

[35]The model is presented since it is the most current published work in the field and does attempt to present the problem as a unified structure in the context of a generalized utility theory.

[36]Lachman.

[37]Ibid., p. 46.

# CHAPTER III

# THE PROSECUTOR AND BARGAINING

In this chapter a model of the prosecutor's behavior is developed. The primary focus is on the influence of variations in the minimum sentence on the charging and bargaining posture of the prosecutor. No specific theory of bargaining is developed. Rather, shifts in the bargaining space because of variations in the minimum sentence are explored.[1] A complete analysis must be postponed until the end of Chapter IV, where a theory of the defendant's behavior, given changes in the minimum sentence, is developed.

## 3.1 A Model of Prosecutor Behavior

The decision to pursue case X vigorously implies that the other cases on the prosecutor's docket will receive less attention, since in most jurisdictions the resources of the prosecutor's office are fully employed. Additional resources can be assigned to case X only by removing resources from other cases.

Let Yo represent years in prison received by defendants other than defendant X. Let Yx represent years spent in prison by defendant X. Let $0 \le c \le 1$ represent an index of crime severity where the more serious the crime the higher the value of c.[2] The prosecuting attorney does not know with certainty the exact Yx that defendant X will receive if convicted of a crime associated with $c = c_x$, but it is assumed that the prosecutor is sufficiently experienced to make an estimate of the value of Yx. It follows then that Yx can be viewed as a random variable with cumulative distribution F and probability density f. For each value of $c_i \in [0,1]$ there exists an $F^i$ and $f^i$. It is further assumed that for each $c_i \in [0,1]$ an $E_{c_i}(Yx) = \int YxdF^i(Yx)$ exists and is greater than or equal to zero.[3] $E_{c_i}(Yx)$ is the expected length of imprisonment upon conviction for crime $c_i$. It is also assumed that $F(0-) = 0$ and that $F(0+)$ can be greater than zero. $F(0-)$ is the left hand limit of $F(Yx)$ at $Yx = 0$ and $F(0+)$ is the right hand limit at $Yx = 0$. In particular, if $F(0+) = 1$ it is assumed that the prosecutor declines to prosecute and that all criminal action against the defendant is terminated. The above condition will be assumed sufficient but not necessary for denial of prosecution by the prosecuting attorney in the case against defendant X.

In general, the shape of F depends upon the crime the prosecutor charges. It will be assumed for sake of simplicity that the shape of F is not dependent upon the charge levied by the prosecutor. If this is reasonable, it is necessary that the range of charges be limited. The superscript from F and f will be dropped and discussion will center around F(Yx) independent of c.

The prosecutor is assumed to maximize an expected utility function dependent upon Yx, p, and c subject to a budget constraint to be developed below. In the above maximization problem, p is a measure of action taken against defendants other than X. In keeping with traditional economics, the expected utility function of the prosecutor will be represented by

$$\int_a^b U(c,p,Yx)dF(Yx),$$

where if $A = \{a_i / P(Yx < a_i) = 0\}$ and $B = \{b_i / P(Yx > b_i) = 0\}$

then $a$ = supremum $\{ai\}$ and $b$ = infimum $\{b_i\}$. Thus a and b set lower and upper bounds, respectively, on the possible length of the prison sentence given the defendant in case X.

It is assumed that the expected utility function U is of the Von Neumann-Morgenstern type[4] and that U is increasing and concave in Yx, c, and p and three times differentiable.[5] The prosecutor is seen as maximizing the above expected utility function by selection of the appropriate c and p while not exceeding the limit of resources made available to him.

Since the prosecutor's resources are limited, it is not possible for him to charge and actively prosecute each defendant for the most serious crime that the defendant is believed to have committed. The case against X can be pursued only by a reduction in resources allocated to other cases. Therefore, any action against X must be analyzed by the prosecutor for its effect upon p.

Let $\frac{R}{T_1} = RP + \frac{RX(c_X)}{T_2}$, where

$\frac{R}{T_1}$ = total variable resources available to the prosecutor, $T_1$ defined below.

RP = value of resources used in prosecution of cases other than X, and

$\frac{RX(c_X)}{T_2}$ = value of resources used in prosecution of defendant x for charge $c_x$, $T_2$ defined below.

Let $T_2$ be a measure of the prosecutor's strength in case X. At the extremes when $T_2 = 1$ and the prosecutor charges defendant X with crime = $c_x$ then the cost of doing so is $RX(c_x)$. At the other extreme when $T_2 = 0$ the prosecutor is incapable of pursuing case X and any attempt to convict defendant X would cost an infinite amount. A reduction in $T_2$ can be viewed as representing an increase in bargaining concessions made by the prosecutor to the defendant. Let $T_2$ be constructed in such a manner that total resources consumed by case X is $\frac{RX(c_x)}{T_2}$. If $c_x$ is obtained via negotiation, its cost is $\frac{RX(c_x)}{T_2^1}$ as opposed to the cost of resolution via trial which is $\frac{RX(c_x)}{T_2^0}$ where $T_2^0 > T_2^1$. Note that $\frac{RX(c_x)}{T_2^1} > \frac{RX(c_x)}{T_2^0}$. This is so since the cost of convicting defendant X of a crime c with severity $c_x$ should increase as his ability to pursue the case decreases because of such things as deterioration or lack of evidence, promises not to pursue certain charges, etc.

As mentioned above, bargaining results in savings to the prosecutor. The resources saved in case X can be used in other cases. This can be viewed as an increase in R. Let this increase in R be represented by $\frac{R}{T_1}$ where a reduction in $T_1$ occurs as a result of bargaining. The new "effective" level of resources then is $R + \Delta R$ where $\Delta R = \frac{-R\Delta T_1}{(T_1)^2}$. $T_1$ is greater than zero and less than one and is equal to one prior to bargaining.

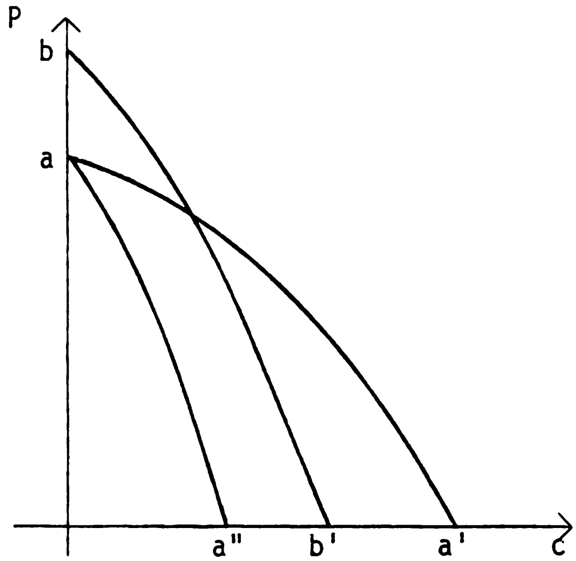



Figure 5.--Changes in the resource constraint curve resulting from bargaining.

In Figure 5 above, the line aa' represents the initial possibilities facing the prosecutor. If only the case against defendant X is pursued then $c_X$ = a' can be obtained via trial. As the charge against defendant X is reduced toward zero (still via trial) the composite charge that can be obtained against other defendants approaches a. Upon the adoption of a bargaining posture the trade-off curve shifts to bb'. This shift can be decomposed into two components. The first is the shift of aa' to aa". This shift stems from the increased cost of achieving a particular conviction given the reduced ability of the prosecutor to vigorously pursue the case against defendant X. The curve bb' is derived by a movement from aa" to bb'. This shift results from the increase in free resources stemming from the bargaining option. This freeing of resources derives not from the actual bargaining with defendant X but rather from the decision to not acquire a conviction by way of jury trial. Thus these freed resources could be used in case X or other cases; thus b' ≥ a". There are two trade-off curves of interest to the prosecutor, namely aa' and bb'. Given b > a if b' were greater than a' it would imply that the prosecutor did not reduce his feasible set by bargaining which is counter to the thinking on bargaining.[6] It would also imply that the utility maximizing prosecutor would always pursue the bargaining option and if c is a normal good actually charge defendant X with a more severe crime under bargaining than via trial. This follows since indifference curve tangent to aa' would by necessity intersect bb' implying that a higher level of total utility could be obtained by moving to bb' which is accomplished by bargaining. The rationale underlying the shifting trade-off constraint has been discussed by most researchers analyzing bargaining in the prosecutor's office. The discussion, however, has never been fully developed except by Lachman.[7] The shifting of the trade-off curve will be more fully developed later in this chapter.

It is further assumed that the cost of prosecution does not decrease with the seriousness of the charge levied against defendant X. That is, $\frac{RX'(c_X)}{T_2} = \frac{dRX(c_X)}{T_2} / dc_X \geq 0$.

The amount of action that can be taken against defendants other than X is dependent upon the level of resources available for their prosecution, which is $RP = \frac{R}{T_1} - \frac{RX(c_X)}{T_2}$. The level of prosecution that can be obtained against defendants other than X given RP is $P = D(RP) = D(\frac{R}{T_1} - \frac{RX(c_X)}{T_2}) = H(c, R, T_1, T_2)$ where

$D' = dD/dRP > 0$, and (14)

$$dP/dc = (dP/dRP)(dRP/d\tfrac{RX}{T_2})(d\tfrac{RX}{T_2}/dc) = D'(-1)\frac{RX'}{T_2} = dH/dc = H_c < 0, \quad (15)$$

$$d(dp/dc)/dT_2 = \frac{D'RX'}{T_2^2} = H_{cT_2} > 0. \quad (16)$$

Condition 14 states that the amount of action the prosecutor can take against defendants other than X increases as the amount of resources for this purpose increases. Condition 15 states that if the level of resources available to the prosecutor is fixed, then increases in the charge against X necessitate a reduction in action taken in other cases. (It is assumed that the prosecutor fully utilizes available resources.) It should also be noted that -dp/dc is the marginal rate of technical substitution.[8] It measures the rate at which p can be substituted for c given a fixed level of resources. Condition 16 states that as $T_2$ (the prosecutor's ability to pursue case X) increases the absolute value of dp/dc falls. That is, the price of c relative to the price of p falls.

The expected utility maximization problem of the prosecutor may be written as: $\underset{c,p}{\text{Max}} \int_a^b U(c,p,Yx)Fy$ subject to

$$\int_a^b (p - D(\frac{R}{T_1} - \frac{RX(c_X)}{T2}))Fy = \int_a^b (p - H(c,R,T_1,T_2))Fy = 0, \text{ where}$$

$Fy = dF(Yx) = f(Yx)dYx.$

The method of Lagrange multipliers[9] will be used for solution of the maximization problem. This results in a new function L to be maximized.

$$L \approx \max_{c,p} \int_a^b U(c,p,Yx)Fy - \int_a^b \lambda(p-H(c,R,T_1,T_2)Fy\,.$$

The first order conditions sufficient for an interior maximum are:[10]

$$dL/dc = \int_a^b UcFy + \int_a^b \lambda HcFy = 0, \tag{17}$$

$$dL/dp = \int_a^b UpFy - \int_a^b \lambda Fy = 0, \tag{18}$$

$$dL/d\lambda = \int_a^b (p-H(c,R,T_1,T_2))Fy = 0. \tag{19}$$

The first order condition in equation 18 states that $\lambda$ is equal to the expected change in total utility because of a small change in P. The first order condition in equation 17 states that at equilibrium the expected change in total utility because of a small change in c must be equal to the expected change in total utility because of a small change in P; a change in P induced by the change in c required if the budget constraint is to be satisfied. The first order condition in equation 19 requires that the utility maximizing prosecutor fully utilize resources made available to him in development of cases against all defendants. The second order condition sufficient to guarantee that equations 17 through 19 represent a maximum and not a minimum is:

$$\begin{Vmatrix} \int_a^b UccFy + \int_a^b \lambda HccFy & \int_a^b UcpFy & \int_a^b HcFy \\ \int_a^b UpcFy & \int_a^b UppFy & -\int_a^b Fy \\ \int_a^b HcFy & -\int_a^b Fy & 0 \end{Vmatrix} > 0. \tag{20}$$

The double bars around 20 indicate that the determinant of 20 should be computed and evaluated using the solutions for c,p and $\lambda$ derived in 17 through 19.

Changes in the minimum sentence, a, can be viewed as a reduction in the level of risk as explained in Appendix A section A.11. In the simplest case risk $r = \theta(a)$ where $\theta' = dr/da < 0$. When the minimum sentence a is increased the cumulative distribution $F(Yx,a,b,z) = F(Yx:r_1)$ shifts to $F(Yx:r_2)$. (a,b, and z are parameters determining the shape of the distribution F.) For small changes in the level of risk $Fr = dF(Yx:r)/dr$ is a measure of the change in the cumulative distribution associated with the random variable Yx induced by a change in the minimum sentence a. This notation is explained in Appendix A section A.12. In like fashion $Fyr = d(f(Yx))/dr$. Fyr is a measure of how the probability density shifts because of a change in the level of risk.

In order to determine the shift in c resulting from a change in total resources, R, or minimum sentence, a, or the measure of relative strength, $T_2$, 1.1 through 1.3 are totally differentiated. This results in:

$$\begin{bmatrix} \int_a^b UccFy + \int_a^b \lambda HccFy & \int_a^b UcpFy & \int_a^b HcFy \\ \int_a^b UpcFy & \int_a^b UppFy & -\int_a^b Fy \\ \int_a^b HcFy & -\int_a^b Fy & 0 \end{bmatrix} \begin{bmatrix} dc \\ dp \\ d\lambda \end{bmatrix} =$$

$$\begin{bmatrix} -\int_a^b UcFyr - \int_a^b HcUpFyr & -\int_a^b UpHcRFy & -\int_a^b UpHcT_1 & -\int_a^b UpHcT_2 \\ -\int_a^b UpFyr + \int_a^b UpFyr & 0 & 0 & 0 \\ 0 & -\int_a^b H_R Fy & -\int_a^b H_{T_1} Fy & -\int_a^b H_{T_2} Fy \end{bmatrix} \begin{bmatrix} df \\ dR \\ dT_1 \\ dT_2 \end{bmatrix} \qquad (21)$$

Using Cramer's rule[11] and solving for dc/dr leads to:

$$dc/dr = \frac{\begin{Vmatrix} -\int_a^b (Uc + \lambda Hc)Fyr & \int_a^b UcpFy & \int_a^b HcFy \\ 0 & \int_a^b UppFy & \int_a^b Fy \\ 0 & -\int_a^b Fy & 0 \end{Vmatrix}}{D}$$

where D is the determinant in equation 2.

Solving for dc/dr =

$$\frac{\int_a^b (Uc + \lambda Hc)Fyr(-\int_a^b Fy \; x - \int_a^b Fy)}{D} + \frac{\int_a^b (Up - \lambda)Fyr( \int_a^b Fy \int_a^b HcFy)}{D}$$

$$+ \frac{\int_a^b (p-H(c,R,T_1,T_2))Fyr( \int_a^b UcpFy - \int_a^b UppFy \int_a^b HcFy)}{D}$$

$$= \frac{\int_a^b (Uc + UpHc)Fyr}{D}$$

since $\int_a^b \lambda Fy = \int_a^b UpFy$ and $\int_a^b Fy = 1$ and $\int_a^b (p-H(c,R,T_1,T_2))Fyr = 0.$[12]

If it is assumed that Up and Uc are concave in Y then the sign of dc/dr is apriori indeterminate since $\int_a^b UpHcFry > 0$ and $\int_a^b UcFry < 0$. But it is possible to develop a set of conditions under which dc/dr < 0 or dc/dr > 0.

## 3.2 Decomposition of dc/dr

As shown above, $dc/dr = \dfrac{\int_a^b (Uc + HcUp)Fry}{D}$ where $Hc = -\int_a^b UcFy / \int_a^b UpFy$ at equilibrium. That $Hc = -\int_a^b UcFy/\int_a^b UpFy$ can be shown by direct substitution into the first order condition equations 17 and 18. Making the substitution for Hc in the equation for dc/dr results in:

$$dc/dr = \frac{\int_a^b UcFry - \dfrac{\int_a^b UcFy \int_a^b Up\, Fyr}{\int_a^b UpFy}}{D} \tag{22}$$

$$= \frac{\int_a^b UcFy \left[\dfrac{\int_a^b UcFry}{\int_a^b UcFy} - \dfrac{\int_a^b UpFyr}{\int_a^b UpFy}\right]}{D}$$

Investigation of the Marginal Rate of Substitution will aid in the interpretation of equation 22. The Marginal Rate of Substitution (MRS) measures the slope of the indifference curve at a given level of total utility. It is a measure of how p can be traded for c or c for p and still maintain the given level of total utility. The MRS is measured under the restriction that all other variables, except the two under consideration, are held constant. Mathematically the MRS can be expressed as

$$-dp/dc = \frac{\int_a^b UcFy}{\int_a^b UpFy} = MRS.$$

Of interest is how the MRS varies with changes in the level of risk. For example, if the MRS decreases when the level of risk falls, i.e., dMRS/dr > 0, then changes in c are compensated for by smaller changes in p (than prior to the reduction in the level of risk). Thus it is easier to compensate for a reduction in c in the sense that a smaller increase in p is now needed if the prosecutor is to remain at the same level of total utility.

The change in the MRS because of a change in the level of risk holding c and p constant is

$$\left.\frac{dMRS}{dr}\right|_{\substack{dp=0\\dc=0}} = d\left\{\frac{\int_a^b UcFy}{\int_a^b UpFy}\right\}\Big|\ dr = \tag{23}$$

$$\frac{\int_a^b UpFy(\int_a^b UcFry) - \int_a^b UcFy(\int_a^b UpFry)}{(\int_a^b UpFy)^2} = \frac{\int_a^b UcFry}{\int_a^b UpFy} - \frac{\int_a^b UcFy}{\int_a^b UpFy}\,\frac{\int_a^b UpFry}{\int_a^b UpFy} =$$

$$\frac{\int_a^b UcFy}{\int_a^b UpFy}\;\frac{\int_a^b UcFry}{\int_a^b UcFy} - \frac{\int_a^b UpFry}{\int_a^b UpFy} = MRS\left\{\frac{\int_a^b UcFry}{\int_a^b UcFy} - \frac{\int_a^b UpFry}{\int_a^b UpFy}\right\} = \frac{D}{\int_a^b UpFy}\,\frac{ddc}{dr}$$

Figure 6, page 58, can aid in the understanding of equation 23. Assume aa' represents the highest level of utility that can be achieved which is consistent with the linear budget constraint ee'. The utility maximization combination of p and c is obtained at the point f. This is the point where aa' is tangent to ee'. bb' and dd' are two possible indifference curves, after the change in the level

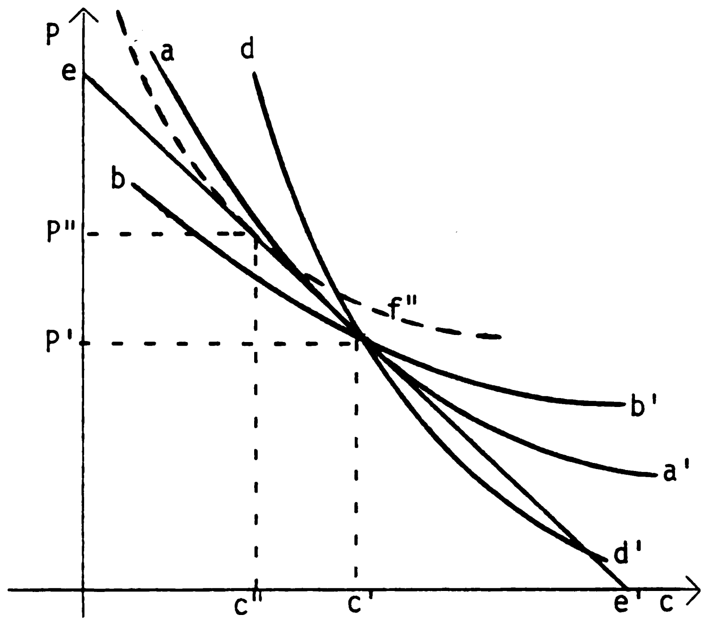



Figure 6.--Determination of the sign of dc/dr.

of risk, which go through f. Since ee' is tangent to aa' at f the slope of ee' is equal to the -MRS of aa' at f. The slope of bb' at f is larger than the slope of ee'. Thus for a reduction in the level of risk dMRS/dr > 0 since the MRS has fallen (i.e., the negative of the slope of bb' measured at the point f has become smaller.) Likewise for dd' the MRS of dd' at f is greater than the MRS of aa' at f. Thus, for a reduction in the level of risk d(MRS)/dr < 0 when measured along dd' at f.

Assume that the relevant indifference curve after the reduction in the level of risk is bb'. Since indifference curves from the same expected utility function cannot cross, it follows that the indifference curve which is just tangent to ee' and is a member of the same family of indifference curves as bb' will generate an optimal solution (f") where oc" < oc' and op" > op'. If dd' were the relevant indifference curve rather than bb' then it is easy to show that the new optimal solution would have oc" > oc' and op" < op'.

Although a linear budget constraint has been used it can be shown that the results hold for nonlinear concave budget constraints except that in this case the inequalities from above (> , <) should be expanded to (≥ , ≤). The easiest way to show this is to assume that the budget constraint is kinked at f. Let the new budget constraint be p'fe' which is still concave. The optimal solution given a reduction in the level of risk and indifference curve bb' is still f and thus oc" = oc' and op" = op'.

### 3.3 Two Cases Where $dc/dr > 0$[13]

At the end of section 3.1 it was mentioned that there existed conditions under which apriori dc/dr could be signed. In this section two sets of assumptions are presented which assert that $dc/dr > 0$. Case 2 presented below, it should be noted, is really a special case of Case 1. These examples are presented as a further aid in developing an intuitive feeling for the condition developed in 3.2 for the signing of dc/dr.

Case 1. In the model formulation the utility function of the prosecutor has, as arguments, two attributes of the case against defendant X. They are c and Yx. If c and Yx are viewed by the prosecutor as substitutes, in the classical meaning of the term, then $Uc_{Yx} < 0$, i.e., the increase in total utility because of a small increase in c decreases as Yx increases. If we assume that $Uc_{Yx}$ decreases at a decreasing rate then $Uc_{YxYx} > 0$. This is symmetrical with the assumption in standard utility theory that marginal utility is positive but decreasing

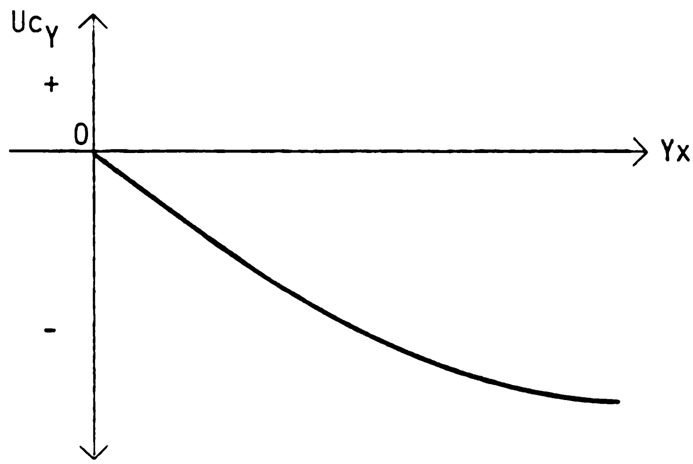



Figure 7.--c and Yx substitutes.

Figure 7 above gives a graphical representation of $Yc_{Yx}$ under these assumptions. The assumption that c and Yx are substitutes does not imply that the prosecutor is risk loving in either c or y.

If $Uc_{YxYx} > 0$ then $\int_a^b UcFyr > 0$, since $\int_a^b UcFyr = UcFr \Big|_a^b - \int_a^b Uc_{Yx}Fr$ $= -\int_a^b Uc_{Yx}Fr = \int_a^b Uc_{YxYx}T(Y)$ for all $Y \in [a,b]$ (see Appendix A.12) is greater than zero as long as $Uc_{YxYx}$ in uniformly signed positive.

Thus if c and Yx are substitutes in the classical sense and Up is concave in Yx, i.e., p and Yx are seen as "complements," then $dc/dr > 0$.

Case 2. This case is a slight modification of Case 1. If $Up_{YxYx} = 0$ and $Uc_{YxYx} > 0$ then $dc/dr > 0$. The interpretation of $Uc_{YxYx}$ was discussed above. $Up_{YxYx} = 0$ means that the marginal utility of p can be affected by Yx but this effect does not vary with the level of Yx. If this is true then $\int_a^b UpFry = 0$ and $dc/dr = \int_a^b UcFyr/D > 0$. This Case 2 is the stochastic analog of the example in footnote 13.

## 3.4 dc/dr in Elasticity Form

The numerator of dc/dr can be written in terms of elasticities.

$$r(dc/dr) = \int_a^b UpFy(MRS) \times \left\{ \frac{rd\int_a^b UcFy \; dr}{\int_a^b UcFy} - \frac{rd\int_a^b UpFy \Big| dr}{\int_a^b UpFy} \right\} \Big/ D$$

$$r(dc/dr) = \frac{MRS(E_{cr} - E_{pr})}{D}$$

where $E_{cr} = \dfrac{\dfrac{rd\int_a^b UcFy}{dr}}{\int_a^b UcFy}$ and $E_{pr} = \dfrac{\dfrac{rd\int_a^b UpFy}{dr}}{\int_a^b UpFy}$

Thus the sign of dc/dr depends upon which marginal utility experiences the greatest percentage change because of the change in the level of risk. If $|E_{cr}| > |E_{pr}|$ then $dc/dr < 0$, otherwise $dc/dr > 0$.

## 3.5 Sign of dc/dR and dp/dR

Using results of the equation in (6) and Cramer's rule then dc/dR, which is the change in charge sought by the prosecutor given a change in resources, may be determined. It is typical for $dc/dR > 0$;[14] in this case c is called a normal good. When $dc/dR < 0$, c is called an inferior good.

$$dc/dR = \frac{\begin{Vmatrix} \int_a^b UpHcRFy & \int_a^b UcpFy & \int_a^b HcFy \\ 0 & \int_a^b UppFy & -\int_a^b Fy \\ \int_a^b H_R Fy & -\int_a^b Fy & 0 \end{Vmatrix}}{D} \tag{24}$$

$$= \frac{\int_a^b UpHcRFy + H_R(\int_a^b UcpFy + \int_a^b HcUppFy)}{D} \tag{25}$$

Given that Ucp  0 and HcR  0 then dc/dR  0. Ucp greater than zero implies that increases in p "enhance" the possession of c. Thus in the traditional sense p and c are viewed as complements.[15] Hc is equal to dp/dc and is a measure of the price of c in terms of p. Thus the absolute value of Hc indicates how much p must be relinquished because of a one unit increase in c. HcR $\geq$ 0 means that the presence of additional resources does not raise the price of c.

If the utility function is separable then Ucp = 0. From inspection of equation 25 it follows that dc/dR is still greater than zero.[16]

The income effect of R on p can be determined in like fashion,

$$dp/dR = \frac{\begin{Vmatrix} \int_a^b (Ucc + UpHcc)Fy & \int_a^b UpHcRFy & \int_a^b HcFy \\ \int_a^b UpcFy & 0 & -1 \\ \int_a^b HcFy & -\int_a^b H_R Fy & 0 \end{Vmatrix}}{D}$$

$$= \frac{\int_a^b UpHcRFy \int_a^b HcFy \; - \; \int_a^b H_R Fy(\int_a^b Ucc + \int_a^b UpHcc)Fy + \int_a^b HcUpcFy}{D}$$

which is greater than zero unambiguously when HcR $\geq$ 0 and Upc $\geq$ 0.

## 3.6 Integration of Rotations in H and Changes in the Level of Risk[17]

## 3.7 Bargaining Without Changes in the Level of Risk

The prosecutor, as developed in this chapter, carries out his function in accordance with the implied rules of constrained utility maximization. The equilibrium levels of c and p are by definition those levels which maximize his utility function given the budget constraint. The budget constraint with which the prosecutor is faced represents the perceived objective trade-off curve between p and c. The trade-off curve is assumed to be a function of such things as relative evidence, cost of prosecution against defendant X, nature of the defendant (number of prior arrests, work history of the defendant, etc.), and resources of prosecutor and defendant. When the prosecutor bargains he relinquishes or mitigates the use of several variables at his disposal for prosecution of defendant X. Thus if the prosecutor, as a result of bargaining, relinquishes the right to use all evidence at his disposal against the defendant the transformation curve will rotate clockwise (see Figure 8). The clockwise rotation implies that after dropping the evidence against the defendant it is more difficult to trade p for c. This would imply that in bargaining the prosecutor really gives up something otherwise the curve would not have rotated. It also implies that he got something since bb' intersects the p axis at a higher level than does aa'.

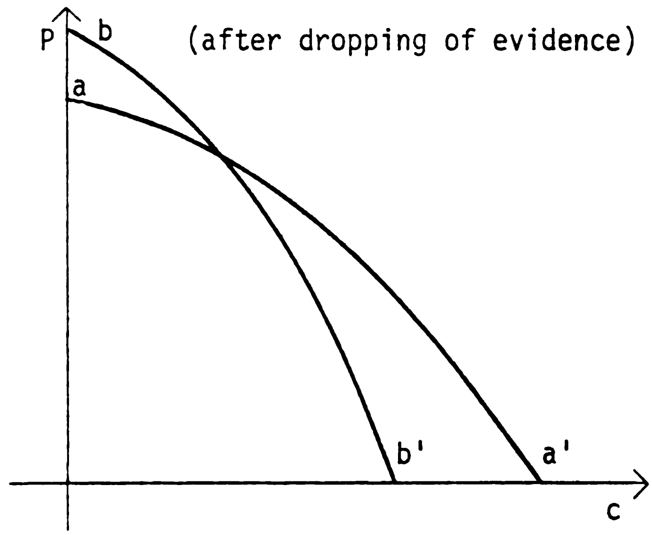



Figure 8.--Two trade-off curves of the prosecutor.

As discussed earlier, the rotation can be separated into two separate effects. They are in fact a price and income effect. Since the prosecutor will, given bargaining, be giving up leverage against the defendant, as described above, the price of c in terms of p rises. Also since bargaining reduces the amount of time needed to handle case X there is a savings in resources which shifts the whole trade-off curve up and to the right (Figures 9 and 10).

Under the trial constraint the prosecutor will select that level of p and c which maximizes his utility. That level of c and p is found at the point where an indifference curve is tangent to the trade-off curve. This represents the highest level of utility that

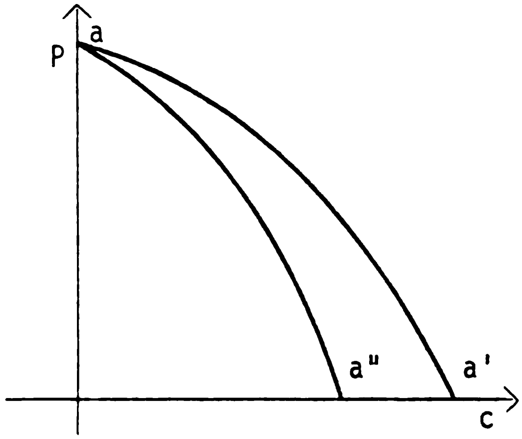



Figure 9.--Shift from aa' to aa" representing the price effect.

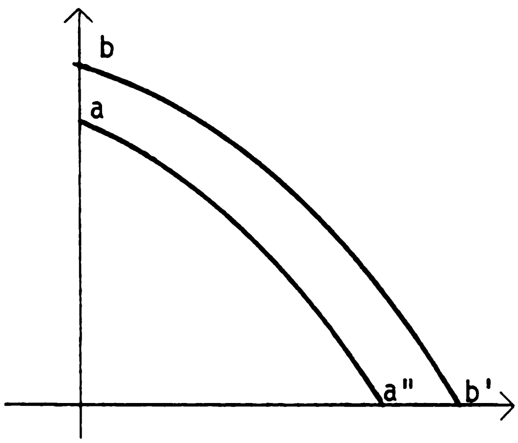



Figure 10.--Shift from aa" to bb' representing the income effect.

can be obtained consistent with the prosecutor's trade-off curve (Figure 11).

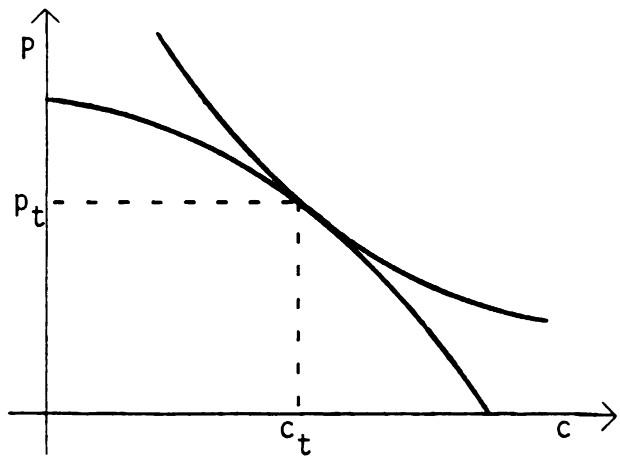



Figure 11.--Level of c and p sought by prosecutor under trial constraint.

As mentioned when the prosecutor considers plea bargaining the cost of prosecuting defendant X is altered. The time needed to search out evidence, prepare briefs, and argue the case are all reduced. This savings will shift the trade-off curve from aa' to dd' (Figure 12).

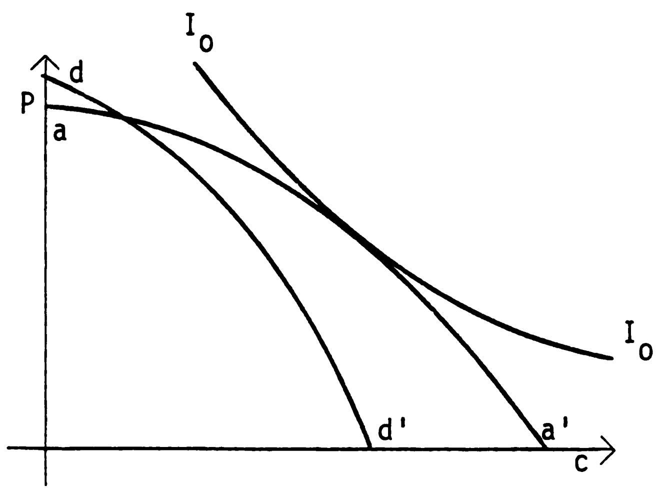



Figure 12.--Preference of trial.

As depicted in Figure 12, it does not always follow that the monetary savings from entering into plea negotiations rather than trial necessarily results in the prosecutor seeking the alternative of plea bargaining. As depicted in Figure 12, entering into plea negotiations by the prosecutor would result in a loss of utility. Thus in this particular instance the prosecutor would seek resolution via trial.

The slope of the trade-off curve is a measure of the prosecutor's ability to trade p for c. As described above, this is a function of the prosecutor's level of evidence against X, etc., and the level of evidence, etc., that the defendant has acquired in his defense. Thus a pledge by the defendant not to use all resources at his disposal will result in a counter clockwise rotation of dd'. If there exists some level of relative resource such that dd' is tangent to $I_0$ than plea bargaining by the prosecutor is consistent with utility maximization as in Figure 13. Thus it is possible through interaction with the defendant to transform a situation like that in Figure 12 to that depicted in Figure 13.

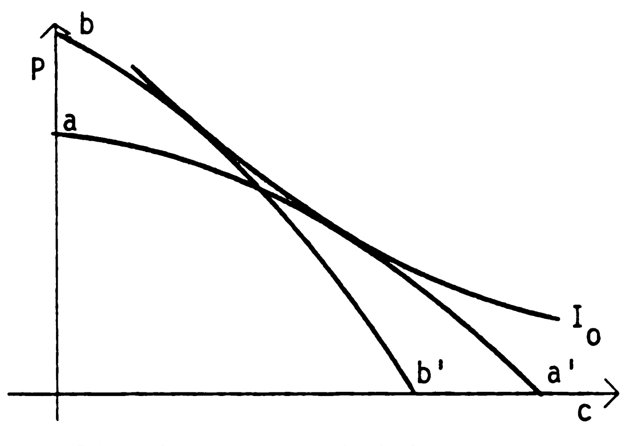



Figure 13.--Prosecutor indifferent between trial and plea negotiating

The initial decision by the prosecutor to bargain rotates the opportunity trade-off curve clockwise from aa' to aa". If this were the only effect of bargaining, the prosecutor would not negotiate since it would leave him on a lower indifference curve as depicted in Figure 14 below.

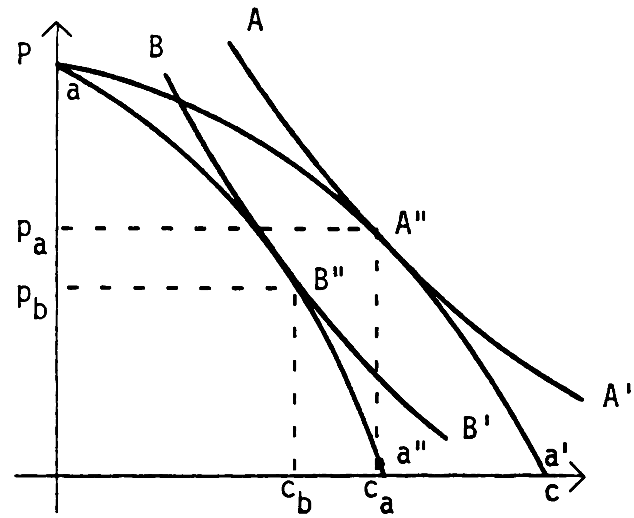



Figure 14.--Shift in c due to increase in the cost of c relative to cost of p.

However, bargaining results in savings to the prosecutor. These savings have the effect of shifting aa" to the right. As the level of savings is increased, aa" shifts further to the right. That aa" which is just tangent to aa' (labeled bb' in Figure 15) represents the <u>minimum</u> savings required from bargaining and <u>minimum</u> charge that the prosecutor will accept if he is to be indifferent between bargaining and trial.

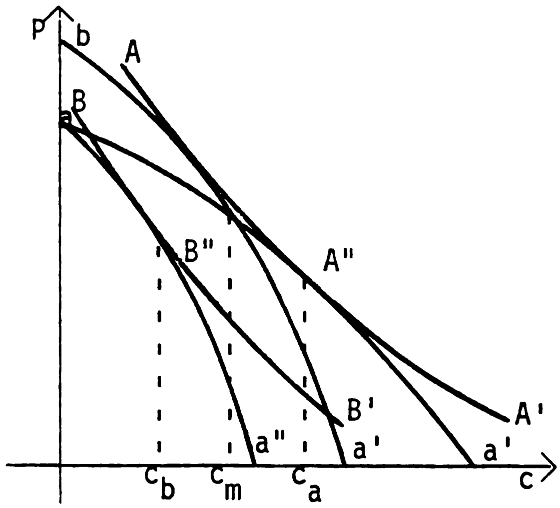



Figure 15.--Separation of shift.

It should be noted that a point like B" in Figure 15 will never occur since the trade-off curve aa" does not exist since the increase in the price of c relative to p and the freeing of resources occur simultaneously. The prosecutor moves directly to a trade-off curve like dd' in Figure 12 or to one like bb' in Figure 15 from aa'. Introduction of aa", however, allows the measurement of the <u>minimum</u> saving needed if the prosecutor is to bargain. In the following section it will be made clear that the substitution effect is a measure of the <u>maximum</u> reduction in c that the prosecutor is willing to consider in exchange for bargaining; and that the Income Effect is a measure of the <u>minimum</u> savings in resources required if the prosecutor is to bargain and not experience a reduction in total utility.

## 3.8 Mathematical Formulation of Shift in Trade-Off Curve

Assume that the act of bargaining raises the cost of c in terms of p to the prosecutor from $(-1/T_2^0)Hc$ to $(-1/T_2^1)Hc$ where $0 < T_2^1 < T_2^0 < 1$. This causes the trade-off curve in Figure 15 to rotate from aa' to aa". This exact movement from $c_a$ in Figure 15 to $c_b$ could be found by total differentiation of the first order conditions with respect to c and $T_2$. This movement can be broken into two parts. The first part $c_a$ to $c_m$ is the own substitution effect and will always be negative. The other part $c_m$ to $c_b$ is the income effect. It can produce either a positive or negative change in c. The own substitution effect is a measure of the maximum reduction in c, from $c_a$ in Figure 15, that the prosecutor can consider and remain indifferent between trial and bargaining. The income effect is a measure of the minimum required savings necessary for the prosecutor to consider bargaining. These effects can be expressed mathematically by solving the system of equations 1.4 for $dc/dT_2$.

$$dc/dT_2 = \frac{\begin{Vmatrix} -\int_a^b UpHcT_2Fy & \int_a^b UcpFy & \int_a^b HcFy \\ 0 & \int_a^b UppFy & -1 \\ -\int_a^b H_{T_2} & -1 & 0 \end{Vmatrix}}{D}$$

$$= \frac{\int_a^b UpHc_{T_2}Fy + \int_a^b H_{T_2}(\int_a^b UcpFy + Hc\int_a^b UppFy)}{D}$$

The income effect is simply dc/dR which was derived earlier. This is multipled by $dR/dT_2 = -H_{T_2}/H_R$ which results in:

$$-\frac{H_{T_2}}{H_R}\left\{\frac{\int_a^b UpHcRFy + H_R(\int_a^b UcpFy + \int_a^b HcUppFy)}{D}\right\}$$

$$= \frac{H_{T_2}}{H_R}\left\{\frac{\int_a^b UpHcRFY \quad - H_{T_2} \quad \int_a^b UcpFy + \int_a^b HcUppFy}{D}\right\} = (dc/dR)(dR/dT_2)$$

The substitution effect is simply $dc/dT_2 + (dc/dR)(dR/dT_2) =$ $\int_a^b UpHcT_2Fy - \frac{H_{T_2}}{H_R}\int_a^b UpHcRFy.$

It appears at first glance that the sign of the substitution term is indeterminate since $HcT_2 > 0$, $H_{T_2} > 0$, $H_R > 0$ and $HcR > 0$ but the substitution term can be rearranged into $\frac{\int_a^b UpFy(HcT_2H_R - H_{T_2}HcR)}{H_R}$ $= H_R\int_a^b UpFy(d(H_{T_2}/H_R)/dc) > 0$ since $d(H_{T_2}/H_T)/dc > 0$. This follows since $H(c,R,T_1,T_2) = D(\frac{R}{T_1} - \frac{RX(c)}{T_2})$ where $H_{T_2} = D'(-1)(\frac{-RX(c)}{T_2^2})$ $= \frac{D'RX(c)}{T_2^2} > 0$. Also $H_R = D'(1/T_1) > 0$ so that

$$\frac{d\left\{\frac{D'RX(c)T_1}{T_2^2D'}\right\}}{dc} = \frac{RX'(c)T_1}{T_2^2} > 0.$$

Thus the substitution term is unambiguously negative as required by the theory. It should be recalled that when $T_2$ is increased the price of c in terms of p falls; thus $-dc/dT_2 = dc/dHc < 0$.

The two parts of the move from $c_a$ to $c_d$ then are:

$$c_a \text{ to } c_m = \frac{\int_a^b UpHcT_2 - \frac{H_{T_2}\int_a^b UpHcRFy}{H_R}}{D}$$

$$\text{and from } c_m \text{ to } c_b = \frac{\frac{-H_{T_2}}{H_R}\left(\int_a^b UpHcRFy + H_R\left(\int_a^b UcpFy + \int_a^b HcUppFy\right)\right)}{D} .$$

## 3.9 Changes in the Level of Risk and Bargaining

It will be shown that given $dc/dr > 0$ that a decrease in the level of risk results in a reduction in $c_m$.[18] (cm being the minimum acceptable plea bargain charge that the prosecutor will accept.)

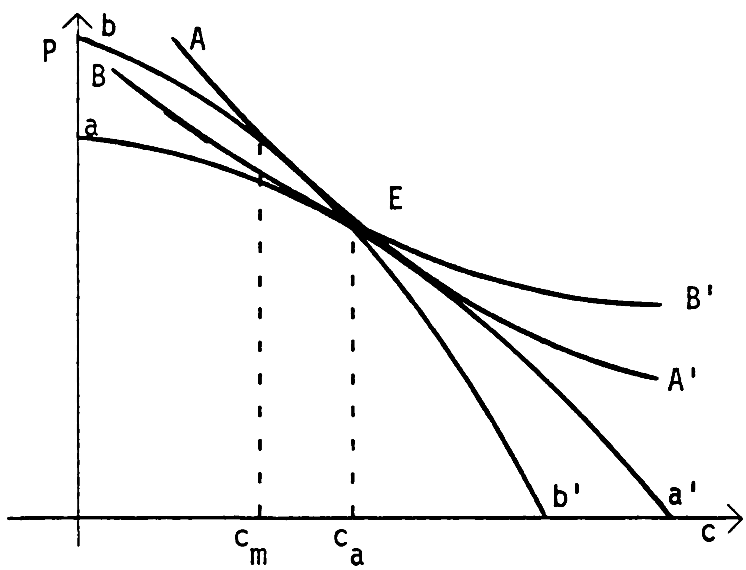



Figure 16.--Shifts in $c_m$ and the level of risk.

In Figure 16 is the trade-off curve given the trial constraint; and the prosecutor's equilibrium position, conditional upon a trial, would be at E. bb' is the trade-off curve given bargaining. With the change in the level of risk the relevant indifference curve is BB'. Since BB' intersect bb' it implies that for a bargain to be preferred to a trial, a smaller savings in resources will be required than would have been required without the change in risk. As long as c is a normal good pulling bb' back toward the origin so that tangency with BB' is possible will result in a reduction of $c_m$.

## 3.10 Initial Position of Prosecutor and Bargaining

Upon the evaluation of the case against defendant X, the prosecutor determines the trade-off possibilities curve. The stronger his case against the defendant the higher will be $T_2$. It will be recalled that the higher $T_2$ the lower the price of c. Let $T_2^0$ be the initial assessment picked by the prosecutor of the strength of his case against X. It is assumed that $T_1^0$ equals 1. Letting $T_1^0 = 1$ implies that charging defendant X with c = 0 does not affect the maximum p regardless of who defendant X happens to be. As $T_2^0$ increases, $c_m$ increases. This can be seen by inspection of Figure 17, on the following page.

One assumption is made concerning the relative position of dd' relative to bb', namely that the slope of dd' exceeds (in absolute terms) that of bb' for any given c. This means that as the prosecutor's trial case against defendant X deteriorates, the cost of c relative to p in bargaining increases. This is identical to the pre-trial case; in the pre-trial case aa' is less negatively sloped than aa". Thus the price of c on aa" exceeds the price of c on aa' for a given c. Also, the price of c as measured by dd' exceeds the price of c as measured by bb' for a given c.

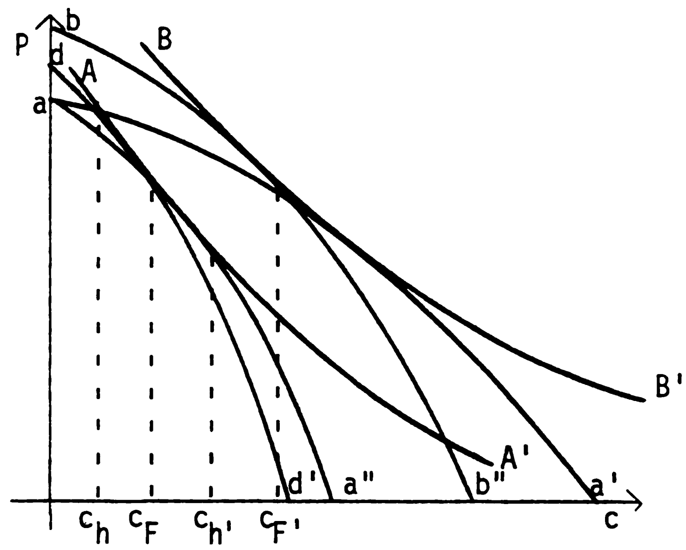



Figure 17.--Shifts in $c_m$.

In Figure 17, $c_h$ is the minimum acceptable plea given the initial bargaining trade-off curve aa', while $c_{h'}$ is the minimum acceptable plea given the initial trade-off curve bb'. It follows from the assumption on the bargaining trade-off curves above that $c_{h'} > c_h$. That this follows can be seen from simple inspection of Figure 17. Assume that bb' and dd' intersect the vertical axis at the same point. Then dd' can be tangent to AA' only by being more steeply sloped than bb'. It follows from the Law of Demand that $c_{h'} > c_h$. If dd' intersects the vertical axis higher up than does bb', then the same argument holds. If dd' intersects the vertical axis below bb' then one of two things (or both) can happen. Assume that the slopes of bb' and dd' are the same. Then as dd' is slid toward the origin to find its point of tangency with AA' the optimal $c_h$ must fall below $c_{h'}$, if c is a normal good. If dd' becomes (at the same or by itself) more steeply sloped then the income effect is reinforced by the own-substitution effect.

## 3.11 Development of the Switch Function

For each $T_2^0$ there exists a unique $c_m$.[19] This relationship can be expressed as $c_m = \theta(T_2^0)$. The above argument suggests that $dc_m/dT_2^0 = \theta' > 0$; that is, when the prosecutor has a stronger pre-bargain case, the minimum c that he will accept in a bargain is larger. This tuple $(T_2^0, c_m)$ is designated as the switch point since $(T_2^0, c_m)$ marks the spot where the prosecutor switches from preferring trial to bargaining[20] (see Figure 18).

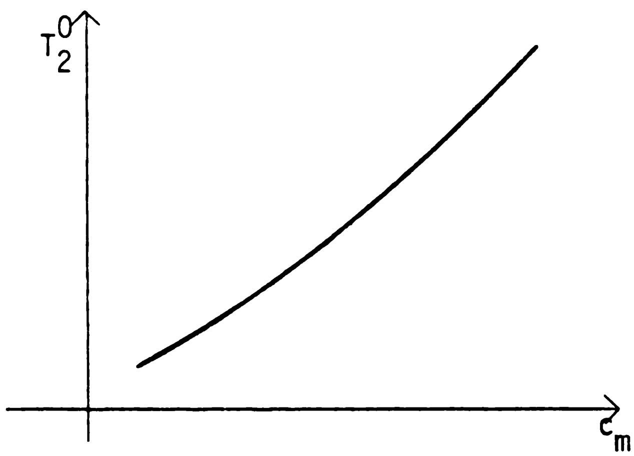



Figure 18.--Switch function.

## 3.12 The Switch Function and Shifts in the Level of Risk

A decrease in the level of risk, given that $dc/dr > 0$, results in a smaller $c_m$ associated with a particular $t_2^0$ to fall. Examination of the following figure will help clarify this point.

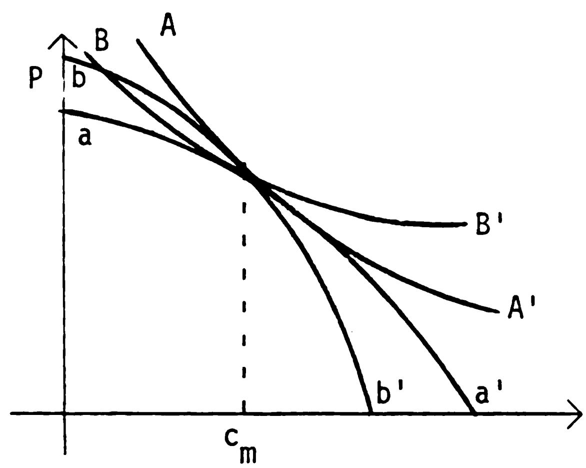



Figure 19.--Shift in the switch function.

In Figure 19 bb' is the trade-off curve given bargaining, while aa' is the trade-off curve corresponding to trial. $c_m$ is the charge when AA' is tangent to bb'. A decrease in the level of risk shifts the indifference curve from AA' to BB'. For bb' to be tangent to BB' requires that the c intercept of bb' move back toward the origin from its position prior to the change in risk. With bb' rotated in this direction the optimal $c_m$ is reduced. Thus a lower $c_m$ is associated with each $T_2^0$ than in the prior case of a higher level of risk. Thus the bargaining space of the prosecutor is increased by a reduction in the level of risk since for any given $T_2^0$ he is willing to accept a lower $c_m$ (see Figure 20).

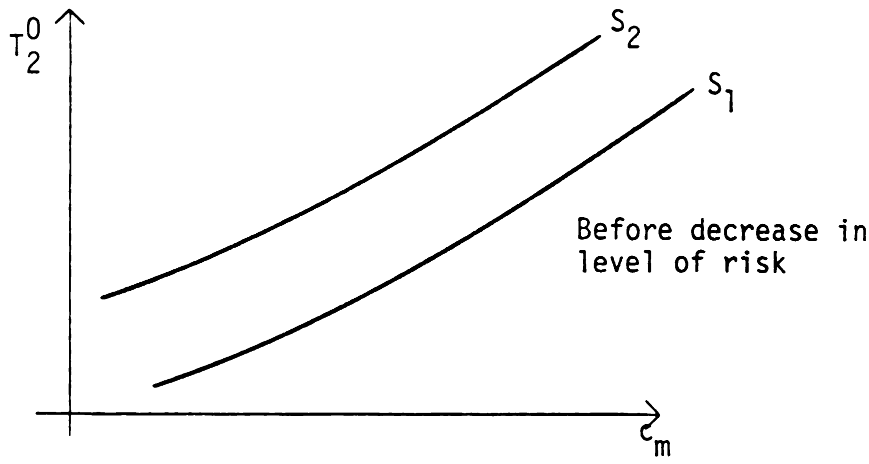



Figure 20.--Shift in the switch function resulting from a decrease in the level of risk.

## 3.13. Two Examples of Expected Utility Maximization by the Prosecutor

Example 1. $\quad U = \int_a^b \ln(1 + c)YxFy + \int_a^b \ln(1 + p)YxFy$

Example 2. $\quad U = \alpha_0 + \int_a^b (\sum_1^3 \alpha_i \ln X_i + \frac{1}{2}\sum_1^3 \sum_1^3 B_{ij} \ln(X_i)\ln(X_j))Fy$

where $X_1 = c$, $X_2 = p$ and $X_3 = Yx$.

Example one is presented because it satisfies all conditions set forth in the theory and is simple enough to allow for explicit solution of the optimal c, dc/da, etc., when using a linear budget constraint. Example two is a form of the translog utility function.[21] The sign of dc/dr will be solved for using the translog utility function above and the budget constraint $H(c,p,R) = 0$.

### 3.13.1 Example 1

The prosecutor maximizes $\int_a^b \ln(1 + c)Y_x^\delta Fy + \int_a^b \ln(1 + p)Y_x^\varepsilon Fy$ subject to $M = w_1c + w_2p$. The Lagrange formulation of the problem is

$$\text{Max } L = \int_a^b \ln(1 + c)Y_x^{\delta}Fy + \int_a^b \ln(1 + p)Y_x^{\varepsilon}Fy - \lambda(M - w_1c - w_2p).$$

The first order conditions are:

E.1.1 $$\int_a^b 1/(1 + c)Y_x^{\delta}Fy = -\lambda w_1$$

E.1.2 $$\int_a^b 1/(1 + p)Y_x^{\varepsilon}Fy = -\lambda w_2$$

E.1.3 $$M - w_1c - w_2p = 0.$$

Solving for c and p using E.1.1-E.1.3 above

$$\int_a^b(1 + p)/(1 + c)\int_a^b Y_x^{\varepsilon}Fy/\int_a^b Y_x^{\delta}Fy = w_1/w_2 =$$

$$(1 + p)/(1 + c) = w_1/w_2 \int_a^b Y_x^{\varepsilon}Fy/\int_a^b Y_x^{\delta}Fy =$$

$$(1 + (M - w_1c)/w_2)/(1 + c) = (w_1/w_2)\int_a^b Y_x^{\varepsilon}Fy/\int_a^b Y_x^{\delta}Fy =$$

$$(w_2 + M) \int_a^b Y_x^{\delta}Fy - w_1\int_a^b Y_x^{\varepsilon}Fy = cw_1\int_a^b(Y_x^{\delta}Fy + Y_x^{\varepsilon}Fy)$$

$$c = ((w_2 + M)\int_a^b Y_x^{\delta}Fy - w_1\int_a^b Y_x^{\varepsilon}Fy)/w_1(\int_a^b Y_x^{\delta}Fy + \int_a^b Y_x^{\varepsilon}Fy)$$

For computational ease, assume that the prosecutor believes that all sentences occur equally between a and b. This assumption of ignorance on the part of the prosecutor as to sentence received by the defendant, except for knowledge of mandated legislative extremes, is not without merit given the nature of actual sentencing practice, which is not under his control.[22]

The cumulative distribution associated with Yx may be written as $F(Yx) = (Yx - a)/(b - a)$ and the probability density $f(Yx) = 1/(b - a)$. What happens under reduction in risk may be analyzed by movement in a (which can be interpreted as the minimum sentence the prosecutor expects the defendant to receive). As illustrated in Figure 21 below, an increase in a represents a reduction in risk. This new distribution F stochastically dominates G.

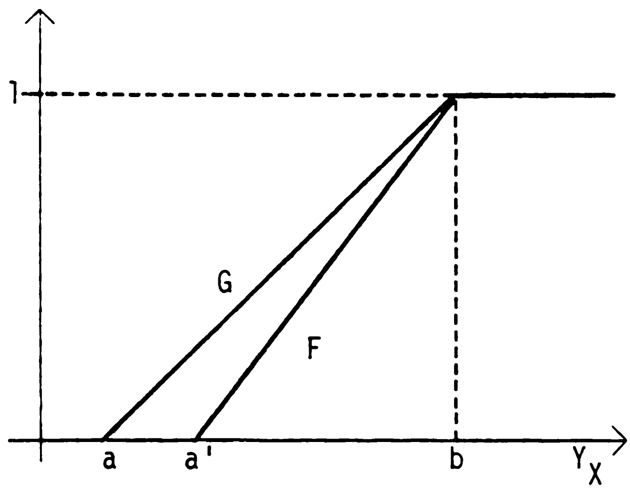



Figure 21.--Cumulative distributions F and G.

This type of shift in a occurred because of the passage of the gun law, which excludes the possibility of sentences under two years in length. Since the law did not affect maximum sentences, it is realistic to not adjust b. It should be noted that if, in fact, the gun law did shift b to b' where b' > b, then F would still dominate G. If b were reduced to b" < b, then in many cases by visual inspection F could be analyzed for $n^{th}$ order dominance of G, n = 2, 3, . . . Or the analysis could be carried out by more sophisticated methods if necessary. For example, second order dominance requires that $\int_a^z ((z - a)/b'' - a) - (z - a)/(b - a))F_z > 0$ for all $z \in (a,b)$. This condition states that the area where G is above F must, at all times, be greater than the area where F is above G (see Figure 22).

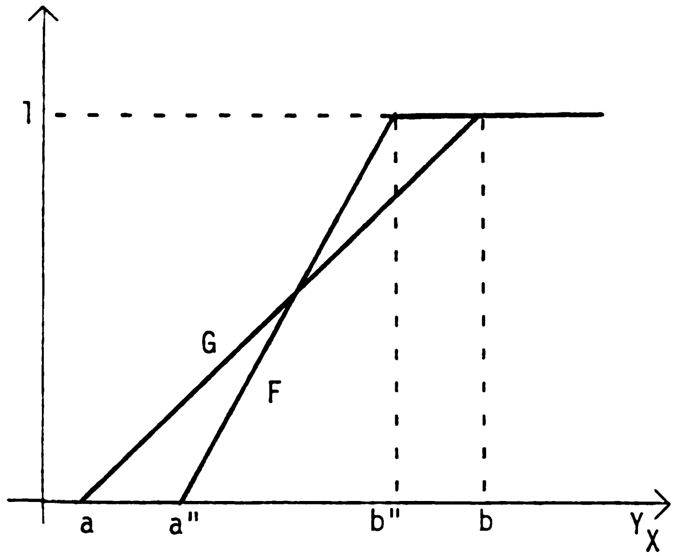



Figure 22.--Case where visual inspection is sufficient to guarantee $\int_a^z S(x)dx > 0$ for all $z \in (a,b)$.

Using then the uniform distribution and substituting in for $\int_a^b YxFy$ and $\int_a^b YxFy$ results in

$$c = \frac{(w_2 + M)(b^{\delta+1} - a^{\delta+1})/((\delta+1)(b-a)) - w_1(b^{\varepsilon+1} - a^{\varepsilon+1})/((b-a)(\varepsilon+1)}{w_1 \ \{(b^{\delta+1} - a^{\delta+1})((\delta+1)(b-a)) + (b^{\varepsilon+1} - a^{\varepsilon+1})/((\varepsilon+1)(b-a))\ \}} =$$

$$\frac{(w_2 + M)(b^{\delta+1} - a^{\delta+1})(\varepsilon + 1) - w_1(b^{\varepsilon+1} - a^{\varepsilon+1})(\delta + 1)}{w_1((b^{\delta+1} - a^{\delta+1})(\varepsilon + 1) + (b^{\varepsilon+1} - a^{\varepsilon+1})(\delta + 1))}$$

as explained previously dc/dr > 0 when dc/da < 0. In order to determine when dc/dr > 0 it is necessary to solve for dc/da first.

$$dc/da = (-a^{\delta}\ (\delta+1)(\varepsilon+1)(w_2+M) + w_1(\varepsilon+1)(\delta+1)a^{\varepsilon})w_1((b^{\delta+1}-a^{\delta+1})$$

$$(\varepsilon+1+(b^{\varepsilon+1}-a^{\varepsilon+1})(\delta+1)) + w_1((\varepsilon+1)(\delta+1)a^{\delta}+(\delta+1)(\varepsilon+1)a^{\varepsilon})$$

$$\frac{((w_2+M)(b^{\delta+1}-a^{\delta+1})(\varepsilon+1) - w_1(b^{\varepsilon+1}-a^{\varepsilon+1})(\delta+1))}{w_1^2((b^{\delta+1}-a^{\delta+1})+(b^{\varepsilon+1}-a^{\varepsilon+1}))^2}$$

dc/da < 0 when

$$\frac{(w_2+M)(b^{\delta+1}-a^{\delta+1})(\varepsilon+1) - w_1(b^{\varepsilon+1}-a^{\varepsilon+1})(\delta+1)}{w_1(((\varepsilon+1)(b^{\delta+1}-a^{\delta+1})+(\delta+1)(b^{\varepsilon+1}-a^{\varepsilon+1})} < \qquad (26)$$

$$\frac{a^{\delta}(w_2+M) - w_1a^{\varepsilon}}{w_1(a^{\delta}+a^{\varepsilon})} \quad \text{or}$$

$$c < \frac{a^{\delta}(w_2+M) - w_1a^{\varepsilon}}{w_1(a^{\delta}+a^{\varepsilon})}$$

To facilitate later work re-write (26) by adding 1 to both sides and obtaining common denominator results in

$$\frac{(b^{\delta+1}-a^{\delta+1})(\varepsilon+1)}{w_1((\varepsilon+1)(b^{\delta+1}-a^{\delta+1})+(b^{\varepsilon+1}-a^{\varepsilon+1})(\delta+1))} < \frac{a^{\delta}}{w_1(a^{\delta}+a^{\varepsilon})} \qquad (27)$$

dc/da < 0 then when equation 26 or 27 holds.

It was shown that dc/dr > 0 when dMRS/dr > 0 (or dMRS/da < 0). In this example MRS = $((1+p)/(1+c))(\int_a^b Y\overset{\delta}{X}Fy/\int_a^b Y\overset{\varepsilon}{X}Fy)$.
The sgn(dMRS/da) = sgn $\{ d(\int_a^b Y\overset{\delta}{X}Fy/\int_a^b Y\overset{\varepsilon}{X}Fy \}$ thus dMRS/da < 0 when $\int_a^b Y\overset{\varepsilon}{X}Fyd \int_a^b Y\overset{\delta}{X}Fy/da - \int_a^b Y\overset{\delta}{X}FydY\overset{\varepsilon}{X}Fy/da < 0 =$

$$- \quad \frac{(b^{\varepsilon+1} - a^{\varepsilon+1})a^{\delta}}{(\varepsilon + 1)} + \frac{(b^{\delta+1} - a^{\delta+1})a^{\varepsilon}}{(\delta + 1)} < 0 \qquad \text{or}$$

$$(b^{\varepsilon+1} - a^{\varepsilon+1})a^{\delta}/(\varepsilon + 1) > (b^{\delta+1} - a^{\delta+1})a^{\varepsilon}/(\delta + 1)$$

Rearranging and adding 1 to both sides, multiplying by $w_1$ and finding a common denominator results in

$$(b^{\delta+1} - a^{\delta+1})(\varepsilon + 1)(w_1((\varepsilon + 1)(b^{\delta+1} - a^{\delta+1}) + (\delta + 1)(b^{\varepsilon+1} - a^{\varepsilon+1}))$$

$$< \frac{a^{\delta}}{w_1(a^{\delta} + a^{\varepsilon})}$$

which is the same as E.1.6.

### 3.13.2 Numerical Example

Let $w_1 = 50$, $w_2 = 1$, $a = 1$, $b = 5$, $\delta = .4$, $\varepsilon = .6$, $M = 100$; then using the formula for U and E.1.4 $c = .3445$, $p = 82.775$ and $U = 8.85$. Assume bargaining raises the price of c from 50 to 60 and ignoring the saving in M results in the new optimal solution $c = .1946$, $p = 88.326$ and $U = 8.7878$.

To find $c_m$ it is necessary to find that M such that utility maximization given $w_1 = 60$ and $M = 100 + \Delta M$ results in $U = 8.85$. By substitution it follows that when $M = 3$ then $c = .2168$, $p = 89.99$ and $U = 8.85$.

### 3.13.3 Decrease in the Level of Risk

Assume that a is now 3 (the original one plus the additional two year mandatory sentence) then given $b = 5$, $w_1 = 50$, $w_2 = 1$, $M = 100$ $\delta = .4$ and $\varepsilon = .6$ then $c = .302$, $p = 84.9$ and $U = 10.66$.

Assume as before that bargaining raises the price of c from 50 to 60 and ignoring the saving in M results in the new optimal solution $c = .157$, $p = 90.587$ and $U = 10.6051$.

To find $c_m$ it is necessary to find that $\Delta M$ such that utility maximization given $w_1 = 60$ and $M = 100 + \Delta M$ results in $U = 10.66$.

By substitution it follows that when $\Delta M = 2$ then $c = .1727$, $p = 91.84$ and $U = 10.66$. Note that after the decrease in the level of risk both $\Delta M$ and $c_m$ have fallen as described in the main body of the chapter.

### 3.13.4 Example 2

The prosecutor maximizes $\alpha_0 + \int_a^b (\sum_1^3 \ln X_i + \frac{1}{2}\sum_1^3 \sum_1^3 \beta_{ij} \ln X_i \ln X_j) Fy$

subject to $H(X_1, X_2, R) = 0$, where $X_1 = c$, $X_2 = p$ and $X_3 = Yx$.

In this example the MRS equals

$$\frac{X_2(\alpha_1 + \beta_{11}\ln X_1 + \beta_{12}\ln X_2 + \int_a^b \beta_{13}\ln X_3 Fy)}{X_1(\alpha_2 + \beta_{22}\ln X_2 + \beta_{21}\ln X_1 + \int_a^b \beta_{23}\ln X_3 Fy)}$$

$$dMRS/dr = \frac{X_2}{X_1}\left\{\frac{\beta_{13}\int_a^b \ln X_3 Fyr(\alpha_2 + \int_a^b \sum_1^3 \beta_{2i}\ln X_i Fy) - (\beta_{23}\int_a^b \ln X_3 Fyr(\alpha_1 + \int_a^b \sum_1^3 \beta_{ij}\ln X_j Fy))}{(\alpha_2 + \int_a^b \sum_1^3 \beta_{2i}\ln X_i Fy)^2}\right\}$$

Now $dMRS/dr > 0$ if

$$\int_a^b \beta_{13} \ln X_3 Fyr(\alpha_2 + \int_a^b \sum_1^3 \beta_{2i} \ln X_i Fy) > \beta_{23} \int_a^b \ln X_3 Fyr(\alpha_1 + \int_a^b \sum_1^3 \beta_{ij} (\ln XjFy)$$

or

$$\frac{(\alpha_2 + \int_a^b \sum_1^3 \beta_{2i} \ln X_i Fy)}{(\alpha_1 + \int_a^b \sum_1^3 \beta_{1i} \ln X_i Fy)} < \frac{\beta_{23}}{\beta_{13}}$$ since $\int_a^b \ln X_3 Fyr < 0$. $\beta_{13} < 0$ since

$c \in [0,1]$. Adding one to both sides and finding a common denominator leads to

$$\frac{\beta_{13}}{\beta_{23} + \beta_{13}} < \frac{H_1 X_1}{\sum_1^3 H_i X_i}$$ where $H_i = \partial H / \partial X_i$.

If $H(c,p,R)$ is homogeneous of degree k then using Euler's theorem[23] leads to a restatement of the condition for $dc/dr > 0$ which is

$$\frac{\beta_{13}}{\beta_{23} + \beta_{13}} > \frac{H_1 X_1}{-RH_R} = dR/dX_1 (X_1 (X_1/R) = 1/\pi X_1, R$$

where $\pi X_1, R$ is the income elasticity of $X_1$.

## Footnotes--Chapter III

[1]This bargaining space has been referred to as the "contract zone." See Richard P. Adelstein, "The Plea Bargain in Theory: A Behavioral Model of the Negotiated Plea," Southern Economic Journal 6 (January 1978): 488-503.

[2]Several measures of crime severity have been developed by criminologists. One such measure is the Sellin and Wolgang measure of crime severity. This measure is discussed in Robert H. Figlio, "The Seriousness of Offenses: An Evaluation by Offenders and Nonoffenders," The Journal of Criminal Law and Criminology 66 (1975): 189-200. This is not the first work in plea bargaining to use a measure of crime severity. See, for example, Brian Forst and Kathleen Brosi, "A Theoretical and Empirical Analysis of the Prosecutor," The Journal of Legal Studies 6 (January 1977): 186.

[3]The limits of integration refer to the maximum and minimum sentences that defendant X can receive. The minimum sentence is a and the maximum sentence is b. A more precise definition is given below.

[4]John V. Neumann and Oskar Morgenstern, Theory of Games and Edonomic Behavior (New York: John Wiley and Sons, 1944), p. 641. Special attention should be paid to chapter one, section three, "The Notion of Utility." A significant implication of their theory for this study is that utility functions can be subjected only to linear transformations. This feature will preserve the sign of cross partials of the utility function.

[5]This assumption is generally regarded as necessary for the type of analysis that will follow. For example, see G. A. Whitemore and M. Findley, eds., Stochastic Dominance (Lexington Books, 1978). Of special interest is the article by J. Hadar and W. R. Russell, "Applications in Economic Theory and Analysis," pp. 293-334.

[6]See, for example, J. Lachman, "An Economic Model of Plea Bargaining in the Criminal Court System" (Ph.D. dissertation, Michigan State University, 1975), p. 40.

[7]Ibid., p. 33.

[8]J. Henderson and Richard E. Quandt, Microeconomic Theory: A Mathematical Approach, 2nd ed. (New York: McGraw-Hill Book Co., 1971), p. 59.

[9]Kevin Lancaster, Mathematical Economics (New York: Macmillan Co., 1968), p. 46.

[10]For an explanation of the mathematical details underlying the first and second order condition which are sufficient for an interior maximum see Taro Yamane, Mathematics for Economists, 2nd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1968), "Maxima and Minima of Functions," chapter 5, pp. 171-230, esp. p. 194. For rules on the correct signing of bordered Hessians in constrained maximization problems see J. Hadar, Mathematical Theory of Economic Behavior (Reading, Pa.: Addison-Wesley, 1971), p. 346.

[11]Blaine Roberts and David Schultz, Modern Mathematics and Economic Analysis (New York: W. W. Norton and Co., 1973), pp. 98-99.

[12]This follows since using integration by parts on $\int_a^b H(c,R)Fry$ results in $\int_a^b H(c,R)Fr/\partial y = 0$ since $\partial H(c,R)/\partial y = 0$.

Proof: To integrate $\int_a^b H(c,R)Fry$ use integration by parts where

$u = H(c,R) \qquad v = Fr$

$du = \partial H(c,R)/\partial y \qquad df = Fry$

Then $\int_a^b udv = H(c,R)Fr \Big|_a^b - \int_a^b \partial H(c,R(Fr/\partial y$, but

$Fr(a) = Fr(b) = 0$ and $H(c,R)/dy = 0$ thus $\int_a^b H(c,R)Fyr = 0$.

[13]It is possible to construct a less general model in which apriori $dc/dr > 0$. Consider the following non-stochastic utility function for the prosecutor. Let $U = U(c,p,Yx)$ where $Yx = \theta(c)$ and $\theta' > 0$ ($\theta' = dYx/dc$). $\theta'$ greater than zero means that an increase in c leads to an increase in the sentence that defendant X will receive. The prosecutor's optimal choice for c and p is derived by solution of $\max U = U(c,p,\theta(c))$ subject to $p = H(c,R)$. The Lagrange formulation of the problem is $\max L = U(c,p,\theta(c)) - \lambda(p - H(c,R))$. The first order conditions for an interior maximum are

1. $Uc + U_{Y_X}\theta' + \lambda Hc = 0$
2. $Up - \lambda = 0$
3. $p - H(c,R) = 0$

Rearranging 1. and 2. produces $\frac{Uc + U_{Y_X}\theta'}{Up} = Hc$. This equation says that the prosecutor will achieve his utility maximum when the rate at which he can substitute p for c and still maintain the same level of total utility $((Uc + U_{Y_X}\theta')/Up)$ is equal to the rate at which p can be technologically substituted for c(Hc). Let $(c^*,p^*)$ be that combination of c and p which satisfies the above equation and hence maximizes the prosecutor's total utility. What is the effect of the mandatory sentence on $(c^*,p^*)$? The mandatory sentence has the effect of reducing $dYx/dc = \theta'$. This is so since the mandatory sentence "compresses" the range that Yx can take on. It is consistent with the empirical data that $\theta(c^*)$ be the same locally both before and after the change in the law.

Prior to the mandatory sentence it was true that $\frac{Uc(c^*,p^*,\theta(c^*)) + U_{Y_X}\theta'}{Up(c^*,p^*,\theta(c^*))} = Hc(c^*,R)$. After introduction of the mandatory sentence

4. $$\frac{Uc(c^*,p^*,\underline{\theta}(c^*)) + U_{Y_X}(c^*,p^*,\underline{\theta}(c^*))\underline{\theta}'(c^*)}{(Up(c^*,p^*,\underline{\theta}(c^*))} < Hc(c^*,R),$$

where $\underline{\theta}' < \theta'$. This says that changes in c now produce smaller changes in Yx. Equation 4. expresses the fact that given $\underline{\theta}'$ the prosecutor is no longer in equilibrium at (c*,p*) since the Marginal Rate of Substitution < Hc(c*,R).

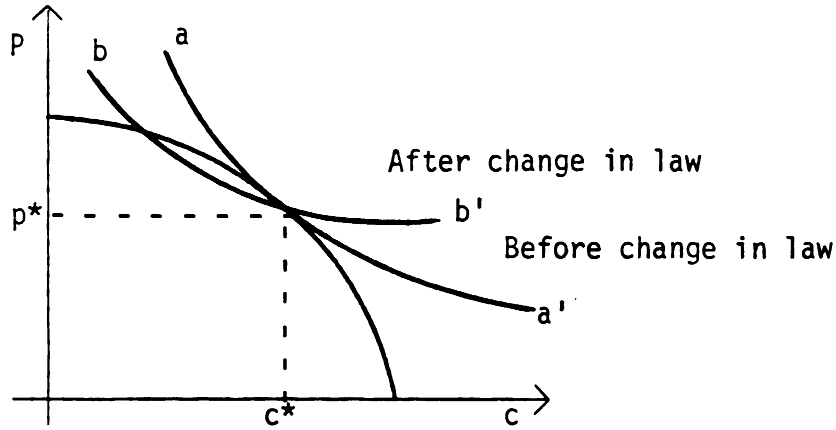



Figure 23.--Shift in c.

The indifference curve bb' is consistent with equation 4. since it goes through (c*,p*) yet has a MRS < Hc(c*,R) at (c*,p*). Since indifference curves from the same family cannot cross it necessarily follows that the new $\underline{c}^*$ < c*. Thus the mandatory sentence results in a reduced charge being sought by the prosecutor. The primary difference between this and the more generalized theory of the text is that in this non-stochastic theory a change in the minimum sentence does not affect Up(c*,p*,θ(c*)) where in the stochastic theory both dU/dc and dU/dp are influenced by the change in risk. The difference is not that in the main body of the dissertation that F(Yx) and Yx are viewed as independent of c. To let F(Yx) = F(Yx,c) and Yx = Yx(c) would greatly complicate the mathematics but would not remove the apriori indeterminancy. It might be enlightening if the reader rather than viewing the mandatory sentence law as a reduction in the level of risk viewed it as an increase in the level of risk that a given change in c would result in the historical change in Yx. Finally the reader should note the similarity between Figure 23 in this footnote and Figure 6 of Chapter 3 for the condition when dc/dr > 0.

[14]William Vickrey, <u>Microstatics</u> (New York: Harcourt, Brace and World, Inc., 1964), p. 58.

[15]Ragnar Frisch, <u>Theory of Production</u> (Chicago: Rand McNally and Co., 1965), p. 60.

[16]If HcR and Ucp = 0 then it also follows that $dc/dR > 0$ since now $dc/dR = H_R Hc \int_a^b UppFy \geq 0$ since $Hc < 0$, $Upp < 0$ and $H_R > 0$. It should be noted that HcR = 0 is not the same thing as saying that the production function is homothetic. Homotheticity implies that $d(dp/dc)/dR = 0 = Hcc(dc/dR) + HcR$. Homotheticity does not necessarily imply HcR = 0. It is possible to derive a sufficient condition for $dc/dR > 0$ if one notes that $dc/dR = \int_a^b dU/dp(dp/dc)/dRFy + \int_a^b (dp/dR(dUp/dc)Fy$ thus $dc/dr > 0$ if $\int_a^b (dUp/dc)Fy > 0$. If an increase in c otherwords increases the marginal utility of p this is sufficient to produce $dc/dR > 0$. (It is assumed that HcR > 0 for the reason noted in the main body of the text.

[17]In this dissertation and the work of Lachman, "An Economic Model of Plea Bargaining in the Criminal Court System" (Ph.D. dissertation, Michigan State University, 1975), it has been assumed that a reduction in the relative strength of the defendant's case to that of the prosecutor results in a counter clockwise rotation of the trade-off curve aa' to bb' in Figure 8. If this rotation is sufficient it could result in the prosecutor opting for plea negotiations rather than trial. This part of the theory is not very controversial and seems plausible.

It is further argued that the level of C sought under bargaining will be less than under trial. This appears reasonable since the analysis is investigating the substitution term of the demand for C. The own substitution is always negative; K. C. Kogiku, Microeconomics Models (New York: Harper and Row, 1971), pp. 20-25. However, the substitution term being negative relies heavily upon the trade-off (budget) constraint being linear; H. W. Folk and J. N. Wolfe, "The Ambiguity of the Substitution Term," Economica 31 (August 1964): 288-93, demonstrate that the substitution term need not be negative when the budget constraint is nonlinear. Their reasoning will be outlined below (see Figure 24, p. 89).

Assume that the trade-off curve dd' shifts to de, where de is assumed to be a straight line for sake of exposition. The price of C is the amount of P given up to obtain a particular amount of C divided by the number of units of C. Thus the price of a quantity oa of C given the initial trade-off curve dd' is dg/oa. To evaluate the effect of shifting from dd' to de on the price of C they look at the relative price of C. Thus,

$$\frac{\text{price of C with constraint de}}{\text{price of C with constraint dd'}} = \frac{dh/oa}{dg/oa} = \frac{dh'}{dg}$$

where $$\frac{dh'}{dg} = \frac{dg + gh}{dg} = 1 + \frac{gh}{dg}$$

This result means that the shift to de implies an unambiguous increase in the price of C. This occurs because de lies entirely below dd'.

To obtain the substitution effect of the increase in price of C the budget line de must be shifted out parallel to itself until tangent to the indifference curve tangent to dd'.

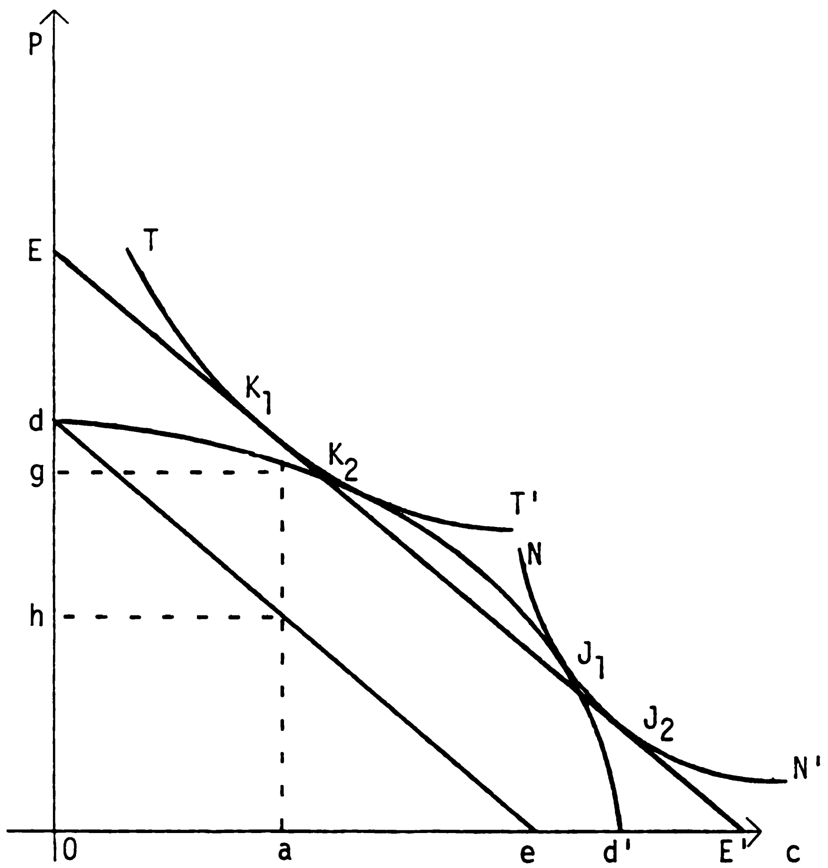



Figure 24.--Possible solutions of utility maximization resulting in ambiguously signed substitution term.

If originally the equilibrium whereat $J_1$, given NN', the substitution effect is positive. If originally equilibrium occurred at $K_2$, given TT', the substitution effect is negative. Thus the substitution term is, a priori, indeterminant.

This result has been contested by R. G. D. Allen and E. J. Mishan, "Is the Substitution Term Ambiguous?" Economica 32 (May 1965): 214-22. Their criticism centers on the definition of price and change in price as used by Folk and Wolf. Folk and Wolfe, they argue, used a concept of average price and average price changes when the substitution term in fact concerns marginal price and marginal price changes -- a point obscured by the fact that given a linear budget constraint the average price and marginal price concepts are identical.

In Figure 25 below, the average price of purchasing OL' units of C is AK/Kl or the slope of AL (sign ignored). The marginal price is measured by the slope of the tangent at L. As can be seen in Figure 26, an increase in the average prices does not always imply an increase in marginal price. They go on to show that the substitution

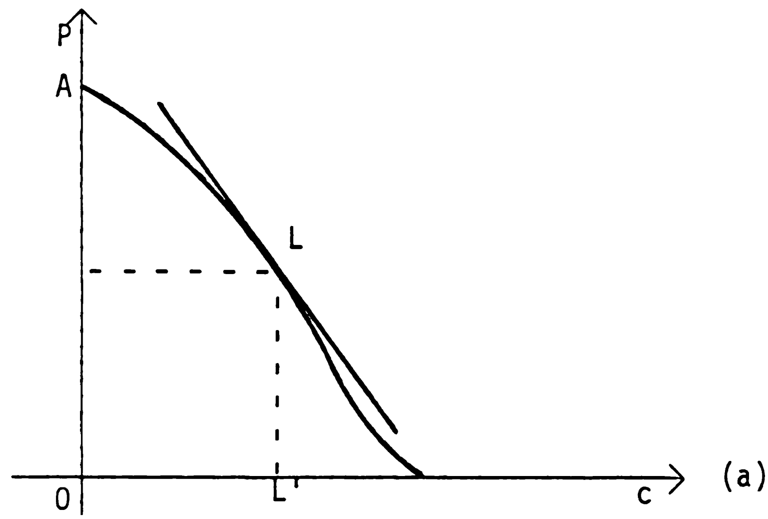



Figure 25.--Transformation curve.

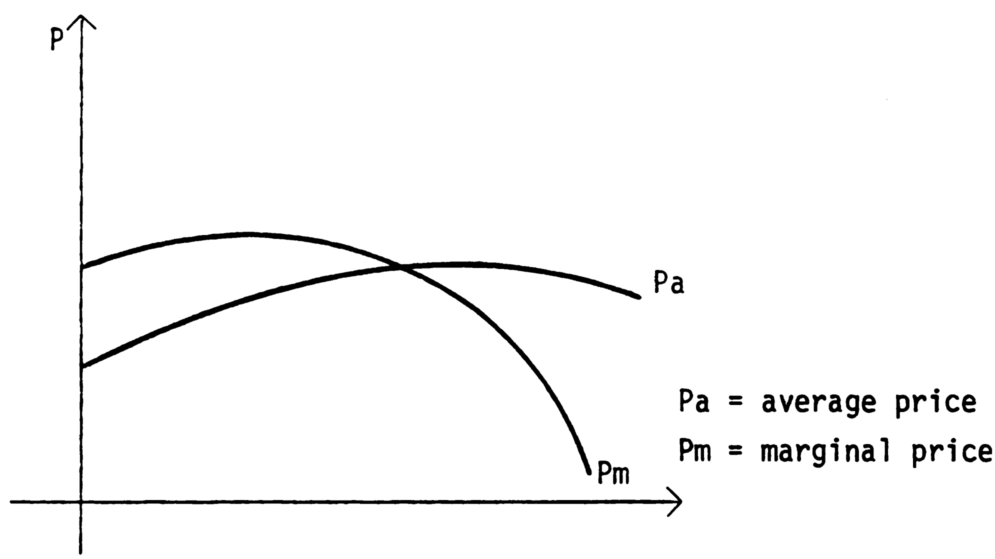



Figure 26.--Influence on average and marginal price.

term is negative when marginal price variations are used. Thus it is possible for the average price to rise but for the marginal price to have falled (when measured at the new equilibrium). Thus to quote Allen and Mishan,

> When a non-linear budget constraint shifts, the analysis is complicated by the fact that the two resultant price changes --in average and in marginal price at the equilibrium position-- may be in the same or opposite directions. It is generally not

possible to say what changes occur in average and marginal price without knowing where the equilibrium positions be on the shifting budget constraints.

[18]A sufficient condition for this to held when dc/dr > 0 is that $\left.\frac{dMRS}{dr}\right|_{dp=0} > 0$. This condition assures that the indifference curve associated with the lower level of risk will be less convex than the indifference curves associated with the higher level of risk. This will assure that a situation like that in Figure 27 cannot happen. If $I_2$ does not intersect $I_1$ to the left of c' then the reduction in risk will result in an increase in the bargaining space.

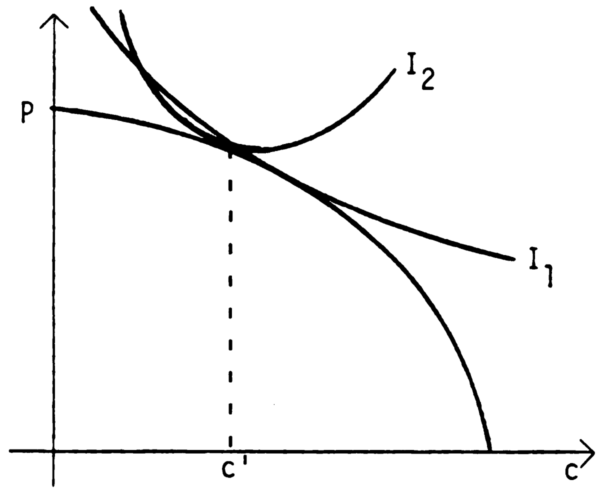



Figure 27.--Case where reduction in the level of risk need not increase the bargaining space.

The indifference curve labeled $I_1$ represents the utility maximization level of c and p before the decrease in the level of risk; $I_2$ represents the same situation after the decrease in the level of risk. It is possible that a rotation of F would result in a lower bargained c' in the state associated with $I_1$ rather than with $I_2$. But if $\left.\frac{dMRS}{dr}\right|_{dp=0} > 0$ for all feasible c and p to the left of c' in Figure 1 then the slope of the indifference curve $I_2$ is less than that of $I_1$ to the left of c'. Working out the math for $\left.\frac{dMRS}{dr}\right|_{dp=0}$ we get

$$\left.\frac{dMRS}{dr}\right|_{dp=0} = \frac{\int_a^b UccFy}{\int_a^b UpFy} - \frac{\int_a^b UcFy \int_a^b UpcFy}{\int_a^b UpFy \int_a^b UpFy} \cdot \frac{dc}{dr} + \frac{1}{\int_a^b UpFy} \int_a^b UcFyr =$$

$$\frac{\int_a^b UcFy \int_a^b UpFyr}{\int_a^b UpFy} = \left.\frac{dMRS}{dc}\right|_{\substack{dp=0\\dr=0}} \frac{dc}{dr} + \left.\frac{dMRS}{dr}\right|_{dc=dp=0}.$$ In the case where

dc/dr > 0 then $\left.\frac{dMRS}{dr}\right|_{dp=0} > 0$. Then the increase in the MRS because of the decrease in c is less in absolute value than the decrease in the MRS because of the reduction in risk.

[19]Lachman, p. 46.

[20]Ibid., p. 47.

[21]D. W. Christensen and L. J. Lau, "Transcendental Logarithmic Utility Functions," American Economic Review 65 (June 1975): 367-83. The direct translog utility function is quadratic in the logarithms of the quantities consumed. It is a local second order approximation to any utility function. It allows expenditure shares to vary with the level of total expenditure and permits a greater variety of substitution patterns among commodities than do utility functions that are homothetic and/or additive.

[22]See Chapter I for a discussion of sentencing practices.

[23]R. G. D. Allen, Macro-Economic Theory: A Mathematical Treatment (New York: St. Martin's Press, 1968), p. 43.

## CHAPTER IV

## THE DEFENDANT AND BARGAINING

### 4.1 The Defendant

The defendant is faced with a basic economic choice between consumption and time in prison or some other disagreeable restriction imposed by the court. The defendant may choose not to negotiate, accept the charge pressed by the prosecuting attorney, and conserve his wealth to be spent by himself or others. Or the defendant may choose to fight the charge and exhaust his entire accumulated wealth attempting to reduce the severity of conviction. Between these two extremes there will be a trade-off between reductions in consumption and decreases in sentence. In the following analysis it is assumed that the defendant possesses some subjective feelings about the exact nature of the trade-off between consumption and sentence reduction.

### 4.2 The Defendant's Utility Function

It is assumed that the defendant has sufficient subjective knowledge about the trade-offs between consumption, conviction for crime C, and the associated imprisonment Y that it is reasonable to represent his preferences over these in a utility function. It is not at all certain that first-time offenders have sufficient understanding and knowledge of the judicial system to make such evaluations with any degree of stability. It will be assumed that between capture and conviction sufficient time will have passed for the defendant to have developed a stable preference relationship. This is one of the functions of defense counsel. Counsel may help the defendant develop a relationship based upon accurate information or upon information that could bias the results. As discussed in Chapter II, lawyers are assigned to indigent defendants by the courts. Many of these lawyers are known for settling out of court. Decisions reached by defendants who have such counsel could be different than those reached be defendants with privately retained defense counsel, given that both defendants have the same resources and the same ambiguous preference structure at the time of arrest. Initially, no consideration will be given to such differences between lawyers. Sufficient information exists within the empirical data, so that this distinction may be tested in later versions of the model.

It is assumed that the defendant maximizes expected utility

a) $\int_{\underline{Z}}^{\overline{Z}} U(V,K,Z)dG(Z)$

subject to b) $K = H(V,W)$ where $W$ = wealth,[1]

$Z = \bar{\bar{Y}} - Y$, $V = 1 - c$, $\underline{z} = \bar{\bar{Y}} - b$ and $\bar{z} = \bar{\bar{Y}} - a$.[2]

Thus the defendant is assumed capable of making precise estimates of the trade-off between V and W. At the same time Z is seen as having a degree of unpredictability.[3] When the defendant enters into a plea bargain for a charge of $c_x$, a charge to be given with certainty, he is still not certain of the actual sentence to be received even when promises have been made by the prosecutor as to the terms of the sentence.[4] This is so since in most jurisdictions, including Michigan's, the prosecutor's office is not officially involved in setting sentences.

## 4.3 Utility Function Concave in V and K

It will be assumed that both V and K are desirable goods and that the defendant's total utility increases with increases in K or V but that it increases at a decreasing rate. Thus

$$\partial U/\partial V > 0, \quad \partial U/\partial V^2 < 0 \quad \text{and} \tag{28}$$

$$\partial U/\partial K > 0, \quad \partial U/\partial K^2 < 0. \tag{29}$$

Equation 28, which states that total utility increases at a decreasing rate as a result of increases in V, implies that total utility decreases because of increases in c at an increasing rate. This is so since

$$\partial U/\partial c = \frac{\partial U}{\partial V}\frac{\partial V}{\partial c} = \frac{\partial U}{\partial V}(-1) = \frac{-\partial U}{\partial V} < 0 \quad \text{and}$$

$$\frac{\partial^2 U}{\partial c^2} = \frac{\partial^2 U}{\partial V^2}(-1)^2 < 0.$$

From Figures 28 and 29 it can be seen that a risk averse behavior of the defendant in V implies a risk averse behavior in terms of c.[5] The same type of symmetry is true of utility functions exhibiting a risk-loving behavior.

## 4.4 The Relationship Between the Tuples (V,Z) and (K,Z)

It is assumed that total utility increases with increases in freedom (Z) and that each unit increase in Z results in successively smaller increases in total utility. Thus $\partial U/\partial Z > 0$ and $\partial^2 U/\partial Z^2 < 0$.

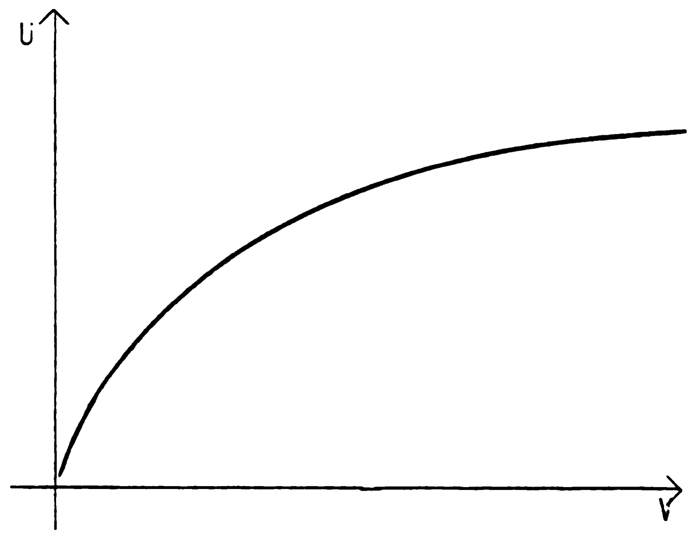



Figure 28.--Change in total utility and V.

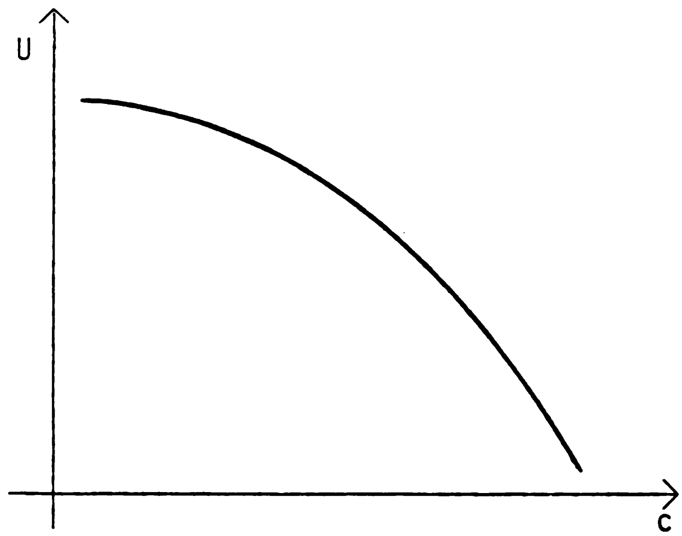



Figure 29.--Change in total utility and c.

These conditions can be seen to be the same as for V and K in equations 4.3.1 and 4.3.2. It can be argued that Z and V should be viewed as substitutes for each other.[6] If this is so then as Z decreases the contribution of a small increase in V (sentence reduction) to increases in utility decreases. As the punishment associated with a given V declines, increases in total utility from increases in V fall since in a sense increases in V (reductions in c) are "buying" less. However, K and Z should be viewed as complements; i.e., as freedom is increased the increment to total utility from a small increase in consumption increases. This seems reasonable if consumption increases in value (in a utility sense) if done at the discretion of the defendant. The likelihood of this happening increases with increased freedom.[7] If the above assumptions about substitutability and complementarity hold, then

$$U_{VZ} < 0 \quad \text{and} \tag{30}$$

$$U_{KZ} > 0. \tag{31}$$

If it is further assumed that the changes occur at decreasing rates

$$U_{VZZ} > 0 \quad \text{and} \tag{32}$$

$$U_{KZZ} < 0. \tag{33}$$

Graphically these two relationships can be depicted as shown in Figures 30 and 31. Of course it is not necessary that these conditions hold. The general results will not depend upon these assumptions concerning complements and substitutes. However, if they do hold, then, given the theory, it will be shown that dV/dr can be signed a priori.[8]

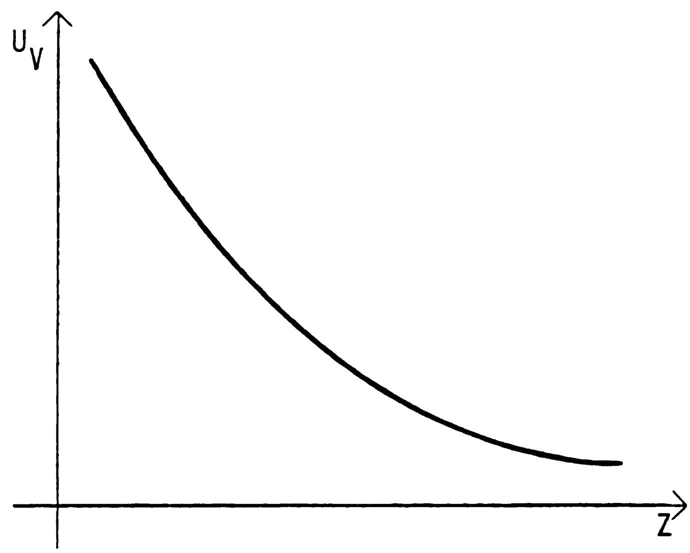



Figure 30.--Effect of Z on the marginal utility of V--"substitute case."

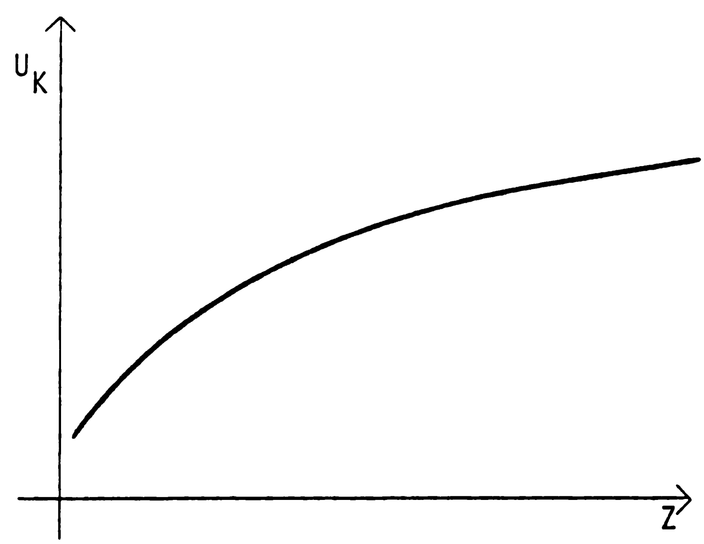



Figure 31.--Effect of Z on the marginal utility of K--"complement case."

## 4.5 The Budget Constraint $K = H(V,W,L_1,L_2)$

The manner in which the defendant can trade wealth and consumption for charge reduction is embodied in $K = H(V,W,L_1,L_2)$. It is assumed that $dK/dV < 0$ and $d^2K/dV^2 < 0$. In other words, increases in charge reduction reduce other forms of consumption and that as efforts are pursued to obtain additional amounts of charge reduction increasing amounts of other consumption must be foregone; i.e., beyond some point charge reduction becomes more difficult to obtain and, if pursued, requires ever accelerating expenditures by the defendant. The trade-off curve then between consumption and charge reduction has the following shape:

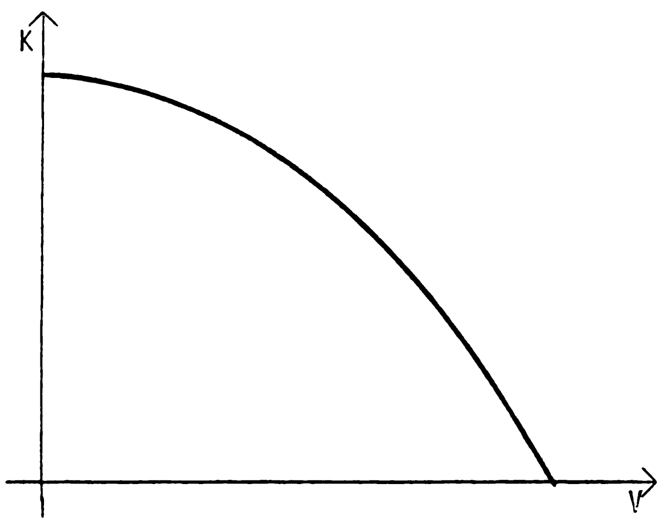



Figure 32.--The defendant's trade-off curve.

$L_1$ and $L_2$ play the same role for the defendant that $T_1$ and $T_2$ did for the prosecutor. Of primary interest is the interpretation of

$L_2$, the measure of the strength of the defendant's case. $L_2$ is constructed and incorporated into $H(\cdot)$ such that $dH_V/dL_2 > 0$; i.e., as the strength of the defendant's case increases, the price of charge reduction in terms of consumption falls. Figure 33 depicts two trade-off curves that differ in terms of initial strength (physical evidence, witnesses, prior arrest record, etc.) where $L_2^1 < L_2^0$.

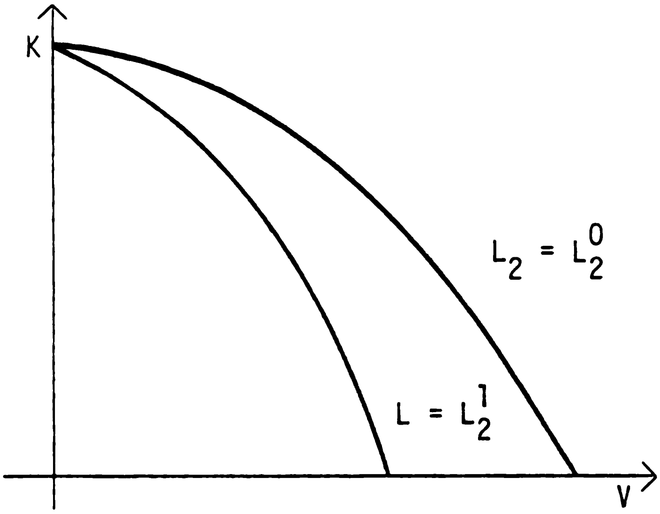



Figure 33.--Effect of variations in $L_2$ on the defendant's trade-off curve.

When the defendant bargains he foregoes the development of a complete defense. In the bargaining process he chooses to not use the entire strength of his case. This is equivalent to $L_2$ falling from $L_2^0$ to say $L_2^1$. If this were the only change that resulted from bargaining the defendant would not engage in it since the bargaining curve lies totally under the trial trade-off curve. But bargaining is cost efficient. The savings stemming from bargaining increase the consumption intercept.

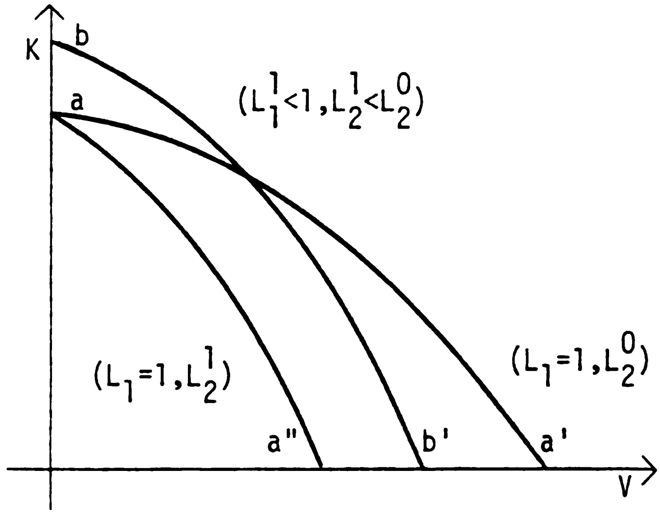



Figure 34.--Total change in trade-off curve.

Thus bargaining results in the trade-off curve shifting aa" to bb'. Reductions in defense effort or strength through bargaining result in a clockwise rotation of aa" while bargaining also results in a shifting out of the trade-off curve from aa" to bb'.

The above leads to the signing of several derivatives in $K = H(V,W,L_1,L_2)$, namely

a) $H_{L_1} < 0$, $L_1$ is a term measuring the effectiveness of W,

b) $H_{L_2} > 0$,

c) $H_{VL_2} > 0$.

Condition a states that as $L_1$ falls, consumption increases. Condition b states that as the defendant's defense strength falls, consumption falls, given that he attempts to acquire conviction for the same crime. Condition c states that as the defense strength falls, the price of charge reduction in terms of consumption loss increases.

## 4.6 The Defendant's Maximization Problem

The defendant desires to maximize expected utility subject to a budget constraint. Mathematically, this is

$$\max_{V,K} \int_{\underline{Z}}^{\overline{Z}} U(V,K,Z)dG(Z) \quad \text{subject to} \tag{34}$$

$$K = H(V,W,L_1,L_2). \tag{35}$$

As in the case of the prosecutor and maximization problem can be rewritten as[9]

$$\max_{V,K} \int_{\underline{Z}}^{\overline{Z}} U(V,K,Z)dG(Z) - \int_{\underline{Z}}^{\overline{Z}} \lambda(K - H(V,W,L_1,L_2))dG(Z) \tag{36}$$

where $\lambda$ is the Lagrangean multiplier.

The first order conditions for an interior maximum are:

$$\int_{\underline{Z}}^{\overline{Z}} U_V G_Z + \int_{\underline{Z}}^{\overline{Z}} \lambda H_V G_Z = 0, \tag{37}$$

$$\int_{\underline{Z}}^{\overline{Z}} U_K G_Z - \int_{\underline{Z}}^{\overline{Z}} \lambda G_Z = 0\ , \tag{38}$$

$$\int_{\underline{Z}}^{\overline{Z}} (-K + H(V,W,L_1,L_2))G_Z = 0. \tag{39}$$

The first order condition in equation 38 states that the Lagrange scalar, $\lambda$, at equilibrium is equal to the expected change in total utility because of a small change in consumption, K. The first order condition in equation 37 states that at equilibrium the expected change in total utility because of a small change in charge reduction, V, must be equal to the expected change in K; this small change in K is equal to the change in K necessitated by the change in V if the budget constraint, equation 35, is to be satisfied. The first order condition in equation 38 states that the defendant at his utility maximum will consume all wealth, W, in obtaining K and V. These conditions, it should be noted, are similar to the first order conditions developed for the prosecutor in section 3.1. The second order condition sufficient to guarantee that 4.5.4-4.5.6 represents an interior maximum and not a minimum is

$$\left\| \begin{matrix} \int_{\underline{Z}}^{\bar{Z}} U_{VV} G_Z + \int_{\underline{Z}}^{\bar{Z}} \lambda H_V & \int_{\underline{Z}}^{\bar{Z}} U_{VK} G_Z & \int_{\underline{Z}}^{\bar{Z}} H_V \\ \int_{\underline{Z}}^{\bar{Z}} U_{KV} G_Z & \int_{\underline{Z}}^{\bar{Z}} U_{KK} G_Z & -1 \\ \int_{\underline{Z}}^{\bar{Z}} H_V & -1 & 0 \end{matrix} \right\| > 0. \tag{40}$$

The double bars around the expressions in 40 indicate that the determinant should be computed and evaluated at V*, K* the optimal V and K as determined in 37-39.[10]

In order to determine the shift in V resulting from a change in total wealth W, minimum sentence a, or the measure of relative strength $L_2$, 37-39 are totally differentiated. This results in:

$$\begin{bmatrix} \int_{\underline{Z}}^{\bar{Z}} U_{VV} G_Z + \int_{\underline{Z}}^{\bar{Z}} \lambda H_{VV} & \int_{\underline{Z}}^{\bar{Z}} U_{VK} G_Z & \int_{\underline{Z}}^{\bar{Z}} H_V \\ \int_{\underline{Z}}^{\bar{Z}} U_{KV} G_Z & \int_{\underline{Z}}^{\bar{Z}} U_{KK} G_Z & -1 \\ \int_{\underline{Z}}^{\bar{Z}} H_V & -1 & 0 \end{bmatrix} \begin{bmatrix} dV \\ dK \\ d\lambda \end{bmatrix} = \tag{41}$$

$$\begin{bmatrix} -\int_{\underline{Z}}^{\bar{Z}}(U_V G_{Zr} + H_V G_{Zr}) & -\int_{\underline{Z}}^{\bar{Z}} U_K H_{VW} G_Z & -\int_{\underline{Z}}^{\bar{Z}} U_{KL_1} G_Z & -\int_{\underline{Z}}^{\bar{Z}} U_K H_{VL_2} G_Z \\ 0 & 0 & 0 & 0 \\ 0 & -\int_{\underline{Z}}^{\bar{Z}} H_W G_Z & -\int_{\underline{Z}}^{\bar{Z}} H_{L_1} G_Z & -\int_{\underline{Z}}^{\bar{Z}} H_{L_2} G_Z \end{bmatrix} \begin{bmatrix} dr \\ dW \\ dL_1 \\ dL_2 \end{bmatrix}$$

Using Cramer's rule[8] and solving for dV/dr results in:

$$\frac{\begin{Vmatrix} \int_{\underline{Z}}^{\bar{Z}}(U_V + U_K H_V) G_{Zr} & \int_{\underline{Z}}^{\bar{Z}} U_{VK} G_Z & \int_{\underline{Z}}^{\bar{Z}} H_V G_Z \\ 0 & \int_{\underline{Z}}^{\bar{Z}} U_{KK} G_Z & -1 \\ 0 & -1 & 0 \end{Vmatrix}}{D}$$

where D is the value of the determinant in equation 40. Solving for dV/dr results in dV/dr =

$$\frac{\int_{\underline{Z}}^{\bar{Z}}(U_V + U_K H_V) G_{Zr}}{D} \; .$$

If $U_{VZZ} > 0$ for all Z and $U_{KZZ} < 0$ for all Z, then $\int_{\underline{Z}}^{\bar{Z}} U_V G_{Zr} > 0$ and $H_V \int_{\underline{Z}}^{\bar{Z}} U_K G_{Zr} < 0$[11] and $dV/dr > 0$ and $dc/dr < 0$ since $-dV/dr = dc/dr$.

## 4.7 The Gun Law and the Level of Risk

It was shown that the passage of the gun law represents a reduction in the level of risk to the prosecutor. It will be shown below that passage of the gun law represents an increase in the level of risk to the defendant. An example is the easiest way to demonstrate this result. Assume that initially the defendant is charged with a crime with severity C and expects the actual sentence received to be uniformly distributed between a and b ($a < b$).

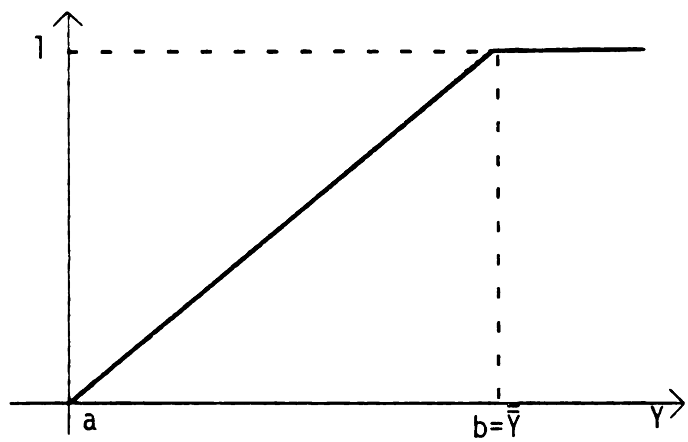



Figure 35.--Cumulative distribution of Y before passage of gun law.

If the variable Y is transformed into sentence reduction Z by letting $Z = \bar{\bar{Y}} - Y$ where $\bar{\bar{Y}}$ is the maximum sentence the defendant expects to receive then Z, like Y, is uniformly distributed but over the interval $(\bar{\bar{Y}} - b, \bar{\bar{Y}} - a)$[12] (see Figure 36). The passage of the gun law increases the minimum sentence that the defendant can receive. It is no longer possible to receive probation or a sentence of less than two years if a gun was used during the commission of the felony. This, in effect, increases the value of a to $a' = a + 2$, where a is the minimum sentence that the defendant could receive before passage of the gun law.

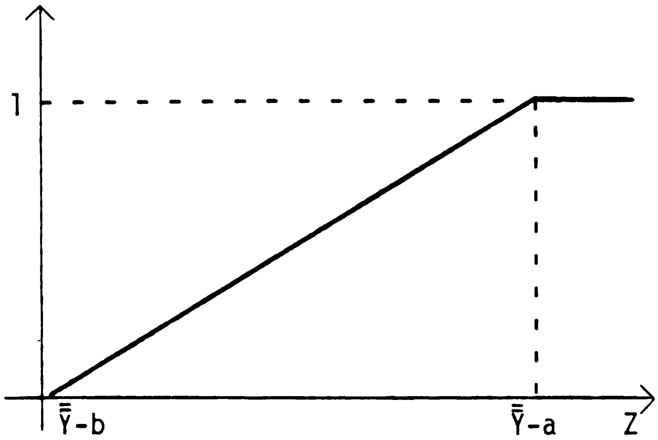



Figure 36.--Cumulative distribution of Z prior to passage of gun law.

As can be seen in Figure 37, the passage of the gun law has resulted in an increase in the level of risk.[13] Thus given assumptions 32 and 38 the passage of the mandatory gun law results in the defendant seeking a reduction in the charge which he would seek via trial. Of course, as in the case of the prosecutor, a generalized condition can be developed for the signing of dV/dr.

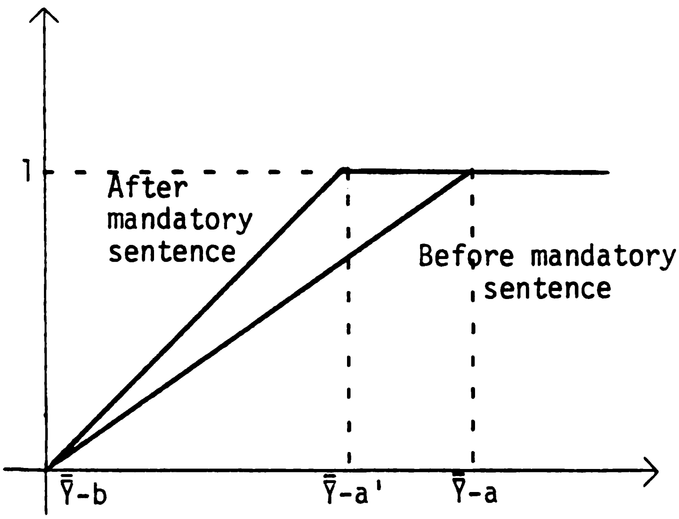



Figure 37.--Cumulative distribution of Z before and after passage of gun law.

## 4.8 The Sign of dV/dr in the General Case

The Marginal Rate of Substitution[14] is

$$\frac{-dK}{dV} = \frac{\int_{\underline{z}}^{\bar{z}} U_V G_Z}{\int_{\underline{z}}^{\bar{z}} U_K G_Z} = -H_V .$$

The change in the MRS given the optimal V and K because of a change in the level of risk is

$$\frac{dMRS}{dr} = \frac{\left\{\frac{d\int_{\underline{z}}^{\bar{z}} U_V G_Z}{dr}\right\} \int_{\underline{z}}^{\bar{z}} U_K G_Z - \left\{\frac{d\int_{\underline{z}}^{\bar{z}} U_K G_Z}{dr}\right\} \int_{\underline{z}}^{\bar{z}} U_V G_Z}{\left(\int_{\underline{z}}^{\bar{z}} U_K G_Z\right)^2} = \tag{42}$$

$$\frac{\left(\int_{\underline{z}}^{\bar{z}} U_V G_{Zr}\right) \int_{\underline{z}}^{\bar{z}} U_K G_Z - \left(\int_{\underline{z}}^{\bar{z}} U_K G_{Zr}\right) \int_{\underline{z}}^{\bar{z}} U_V G_Z}{\left(\int_{\underline{z}}^{\bar{z}} U_K G_Z\right)^2} = \tag{43}$$

$$\frac{\int_{\underline{z}}^{\bar{z}} U_V G_{Zr}}{\int_{\underline{z}}^{\bar{z}} U_K G_Z} - \frac{\int_{\underline{z}}^{\bar{z}} U_K G_{Zr}}{\int_{\underline{z}}^{\bar{z}} U_K G_Z} \; \frac{\int_{\underline{z}}^{\bar{z}} U_V G_Z}{\int_{\underline{z}}^{\bar{z}} U_K G_Z} = \tag{44}$$

$$\frac{\int_{\underline{z}}^{\bar{z}} U_V G_{Zr} + \int_{\underline{z}}^{\bar{z}} U_K G_{Zr} H_V}{\int_{\underline{z}}^{\bar{z}} U_K G_Z} \quad \text{i.e.,} \tag{45}$$

$$\left.\frac{dMRS}{dr}\right|_{dK=dV=0} = \frac{1}{\int_{\underline{z}}^{\bar{z}} U_K G_Z} \cdot \left[ D \cdot \frac{dV}{dr} \right].$$

Thus the sign of dV/dr is the same as the sign of $\left.\frac{dMRS}{dr}\right|_{dK=dV=0}$.

As in the case of the prosecutor, this condition for the signing of dV/dr is readily amenable to graphical representation:

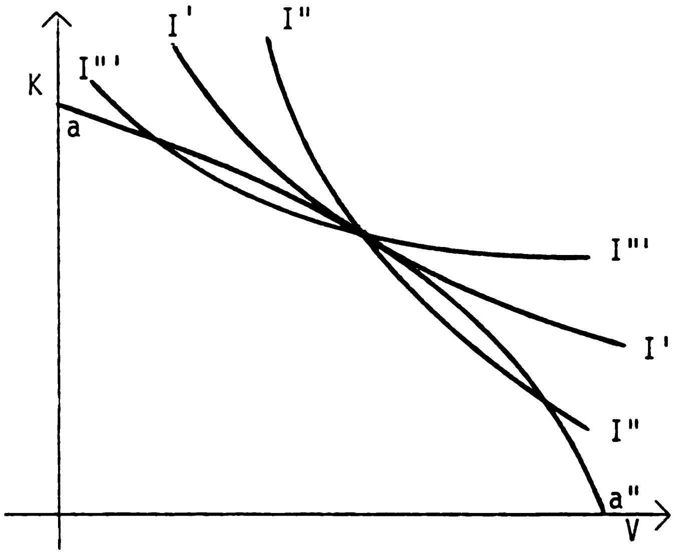



Figure 38.--The signing of dV/dr.

In Figure 38, aa' = H(V,W) is the trade-off curve for V given the trial constraint. I'I' is the indifference curve which is tangent to aa' giving the optimal solution V*,K* derived from equations 37-39. I"I" and I"'I"' are two possible indifference curves given the mandatory minimum sentence law that go through the point V*,K*. If I"I" is the relevant indifference curve and not I"'I"', given the increase in the level of risk, then dMRS/dr > 0, since the increase in the level of risk has resulted in a higher MRS at the point V*,K*. This higher MRS implies that dc/dr < 0 or dV/dr > 0. Exactly the opposite is true if I"'I"' is the relevant indifference curve after the change in risk.

## 4.9 Components of $dV/DL_2$

As in the case of the prosecutor, it is possible to disaggregate $dV/dr$ into income and substitution terms. The income effect is

$$\frac{dV}{dW} = \frac{\begin{Vmatrix} -\int_{\underline{Z}}^{\overline{Z}} U_{K} H_{VW} G_Z & \int_{\underline{Z}}^{\overline{Z}} U_{VK} G_Z & \int_{\underline{Z}}^{\overline{Z}} H_V \\ 0 & \int_{\underline{Z}}^{\overline{Z}} U_{KK} G_Z & \int_{\underline{Z}}^{\overline{Z}} G_Z \\ \int_{\underline{Z}}^{\overline{Z}} H_W G_Z & -1 & 0 \end{Vmatrix}}{D} = \tag{46}$$

$$\frac{\int_{\underline{Z}}^{\overline{Z}} U_K H_{VW} G_Z + \int_{\underline{Z}}^{\overline{Z}} H_W G_Z \left(\int_{\underline{Z}}^{\overline{Z}} U_{VK} G_Z + \int_{\underline{Z}}^{\overline{Z}} U_{KK} G_Z \int_{\underline{Z}}^{\overline{Z}} H_V\right)}{D}$$

The substitution effect is $\frac{dV}{dL_2} + \frac{dV}{dW}\frac{dW}{dL_2}$ where $\frac{dW}{dL_2}$ is $\frac{-H_{L_2}}{H_W}$.[15]

Thus determination of $dV/dL_2$ will allow the evaluation of the substitution term.

$$\frac{dV}{dL_2} = \frac{\begin{Vmatrix} -\left(\int_{\underline{Z}}^{\overline{Z}} U_K H_{VL_2} G_Z + \int_{\underline{Z}}^{\overline{Z}} H_{VL_2} G_Z\right) & \int_{\underline{Z}}^{\overline{Z}} U_{VK} G_Z & \int_{\underline{Z}}^{\overline{Z}} H_V G_Z \\ 0 & \int_{\underline{Z}}^{\overline{Z}} U_{KK} G_Z & -1 \\ \int_{\underline{Z}}^{\overline{Z}} H_W G_Z & -1 & 0 \end{Vmatrix}}{D} = \tag{47}$$

$$\frac{\int_{\underline{z}}^{\bar{z}} U_K H_{VL_2} G_Z + \int_{\underline{z}}^{\bar{z}} H_{VL_2} G_Z + \int_{\underline{z}}^{\bar{z}} H_{L_2} G_Z \left( \int_{\underline{z}}^{\bar{z}} U_{VK} G_Z + \int_{\underline{z}}^{\bar{z}} H_V G_Z \int_{\underline{z}}^{\bar{z}} U_{KK} G_Z \right)}{D} .$$

The substitution term is therefore:

$$\frac{\int_{\underline{z}}^{\bar{z}} U_K H_{VL_2} G_Z + \int_{\underline{z}}^{\bar{z}} H_{VL_2} G_Z + \int_{\underline{z}}^{\bar{z}} H_{L_2} G_Z \left( \int_{\underline{z}}^{\bar{z}} U_{VK} H_Z + \int_{\underline{z}}^{\bar{z}} H_V G_Z \int_{\underline{z}}^{\bar{z}} U_{KK} G_Z \right)}{D} + \tag{48}$$

$$\frac{-\int_{\underline{z}}^{\bar{z}} H_{L_2}}{\int_{\underline{z}}^{\bar{z}} H_W G_Z} \quad \frac{G_Z \left( \int_{\underline{z}}^{\bar{z}} U_K H_{VW} G_Z \right) - \int_{\underline{z}}^{\bar{z}} H_{L_2} G_Z \left( \int_{\underline{z}}^{\bar{z}} U_{VK} G_Z + \int_{\underline{z}}^{\bar{z}} U_{KK} G_Z \int_{\underline{z}}^{\bar{z}} H_V G_Z \right)}{D}$$

$$= \int_{\underline{z}}^{\bar{z}} U_K G_Z \left( H_{VL_2} - \frac{H_{L_2} H_{WV}}{H_W} \right) = \int_{\underline{z}}^{\bar{z}} U_K G_Z \frac{d \left\{ \frac{H_{L_2}}{H_W} \right\}}{dV} .$$

Thus $\left. \frac{dV}{dL_2} \right|_{dU=0} > 0$ since $\frac{d \left\{ \frac{H_{L_2}}{H_W} \right\}}{d V} > 0.$

The above result depends upon $\frac{d \left\{ \frac{H_{L_2}}{H_W} \right\}}{d V} = \frac{d \left\{ -\frac{dW}{dL_2} \right\}}{dV} > 0.$ This states

that as the level of charge reduction sought is increased say from $V_1$ to $V_2$ ($V_2 > V_1$), it is more difficult to maintain the same level of utility for a given fall in defense strength. It is more difficult in the sense that a larger increase in W (used for the defense) is required to maintain the given level of utility given a fall in

$L_2$ than is required to maintain the level of utility given $V_1$ and the same fall in $L_2$. The movement from $V_a$ to $V_m$ is the substitution term. $V_m = (1 - c_m)$ represents the maximum charge that the defendant will accept given plea bargaining. The movement from $V_m$ to $V_b$ is the income effect (see Figure 39).

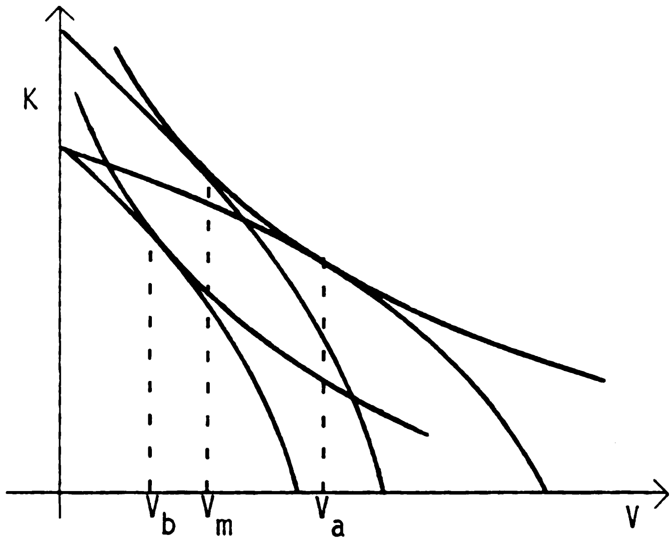



Figure 39.--Shifts in V given a decrease in $L_2$.

## 4.10 Bargaining and Changes in the Level of Risk

Assume that prior to a change in the level of risk that the situation of the defendant is as pictured in Figure 40, on the following page. As depicted in the figure, bargaining would result in the defendant's acceptance of a higher charge than would have been accepted if resolution of the case were via trial.

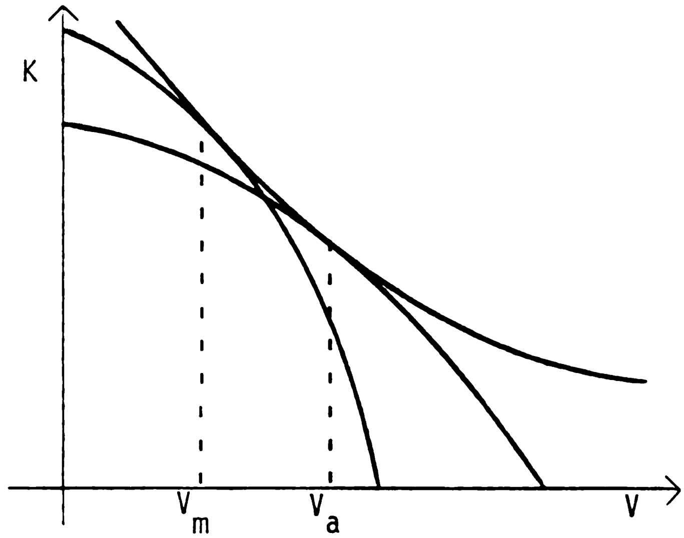



Figure 40.--Defendant indifferent between trial and bargaining.

With the introduction of a change in the level of risk $V_m$ is no longer the optimal minimum plea bargained V. Rather $V_m'$ is the minimum V that is acceptable via plea bargaining. $V_m''$ is greater than $V_m'$ if dV/dr > 0 given G' in Figure 41 and V is a normal good.

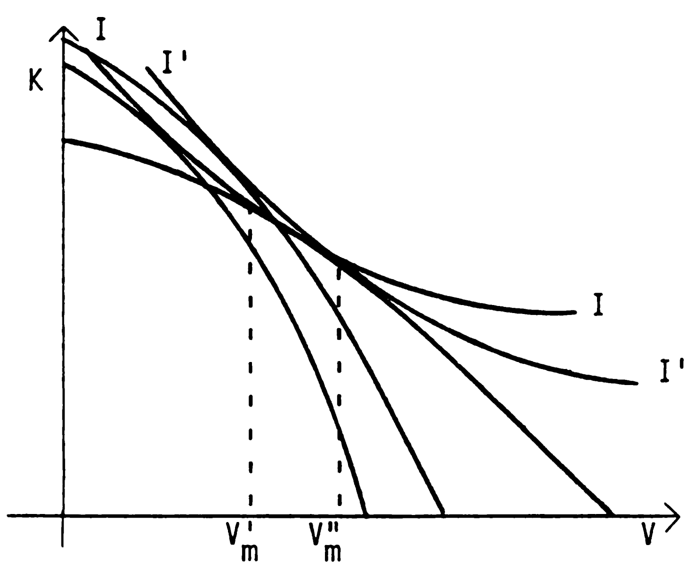



Figure 41.--$V_m$ and shifts in the level of risk.

## 4.11 The Switch Function and Shifts in the Level of Risk

As in the case of the prosecutor, for each $L_2^0$ there exists a unique $V_m$. This relationship can be expressed as $V_m = \rho(L_2^0)$. As in the case of the prosecutor, $\rho' > 0$; i.e., as the initial pre-bargaining strength of the defendant increases so does the level of $V_m$ sought. The tuple $(V_m, L_2^0)$ will be as before labeled the switch point.

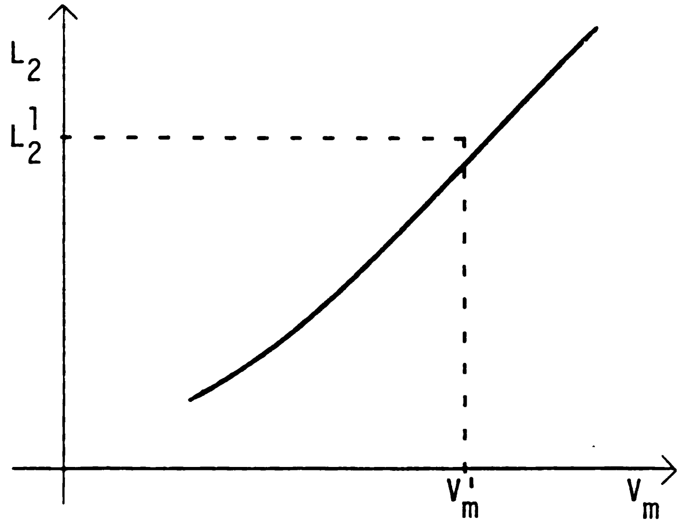



Figure 42.--The defendant's switch function.

For initial strength $L_2^1$ in Figure 42 $V_m^1$ represents the minimum $V_m$ that the defendant will accept via bargaining.

The demonstration that $\rho' > 0$ is similar to that used in the case of the prosecutor to show that $\theta' > 0$. In Figure 43 aa' represents the initial pre-trial trade-off curve for the defendant, where E is the defendant's equilibrium given aa'. Given bargaining,

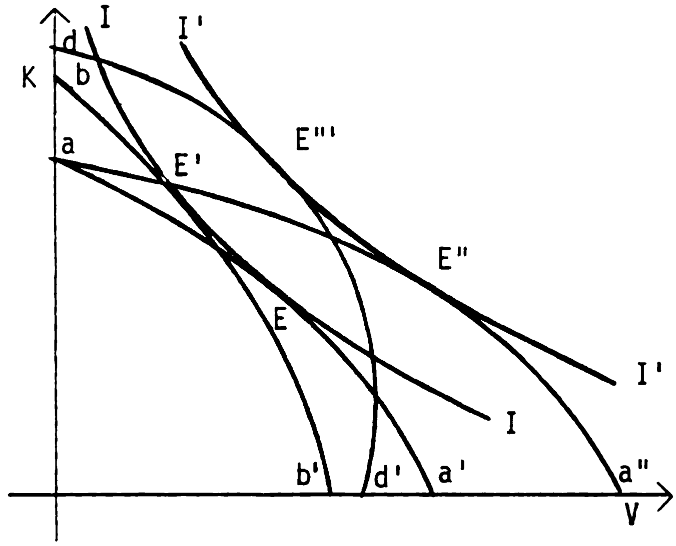



Figure 43.--$L_2^0$ and $V_m$.

the trade-off curve which is just tangent to II, and thus leaves the defendant indifferent between bargaining and trial is bb'. The equilibrium point E' implies a lower V since the generation of E' is equivalent to the own-substitution term. Assume that the strength of the defendant's pre-trial case increases. This is represented in Figure 43 by aa' shifting to aa". The equilibrium is now E"; E" implies a higher V than does E or E'. This follows since the shift in aa' to aa" is equivalent to a decrease in the price of V (this assumes V is not a Giffen good). The basic arguments developed in Section 3.10 in the case of the prosecutor can be used to demonstrate that the V associated with the equilibrium D"' is greater than the V associated with E'. In the case of the defendant, the arguments of Section 3.10 result in dd' being less steeply sloped than bb' for any given c. This is identical to the assumption made concerning the prosecutor and bargaining. Briefly, that assumption stated that since aa" is less steeply sloped than aa' then dd' is less steeply sloped than bb'. In Section 3.10 this assumption is explained more completely.

In Figure 44 below, G is the trade-off curve of the defendant; it is based upon an initial level of power $L_2^0$. An increase in risk, given dV/dr > 0 results in a higher level of $V_m$ being sought by the defendant. The implication of this is that an increase in the level of risk results in the defendant's switch function shifting out and to the right.

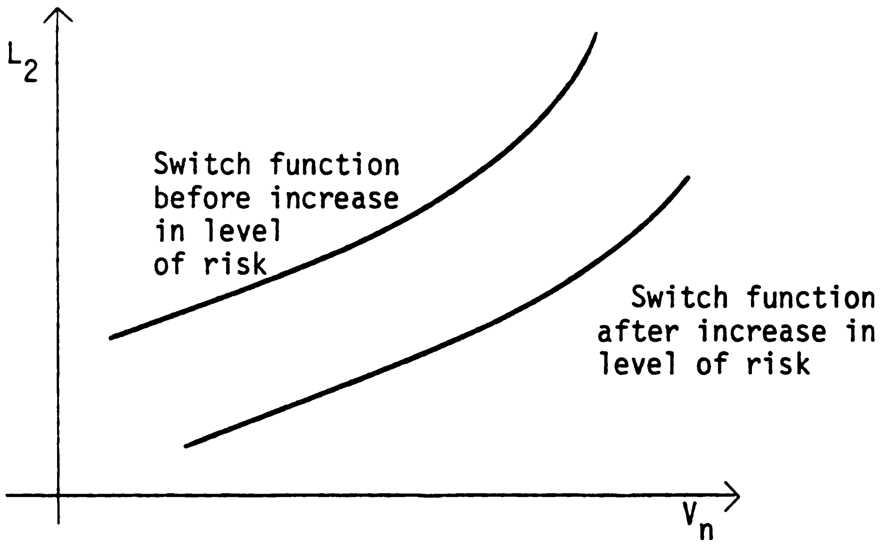



Figure 44.--The switch function and changes in the level of risk.

Thus an increase in the level of risk results in a reduction in the bargaining space of the defendant since it is the $L_2$,V space to the right of the switch curve which represents the feasible bargaining combinations of defense strength and V.

## 4.12 The Likelihood of Bargaining

### 4.13 The Bargaining Framework

It will be shown below that the passage of the gun law, given $dc/dr > 0$ and $dV/dr > 0$ reduces the probability that the agreed-upon charge $c_{pd}$ is greater than or equal to Q, some fixed number between zero and one. The complement to this is that the passage of the gun law in bargained cases increases the probability that $c_{pd}$ is less than Q. Thus the passage of the gun law increases the likelihood of observing less severe charges being pressed by the prosecutor, to which the defendant will plead guilty, <u>ceteris paribus</u>. In the Edgeworth box below the switch functions of both the prosecutor and defendant have been drawn.

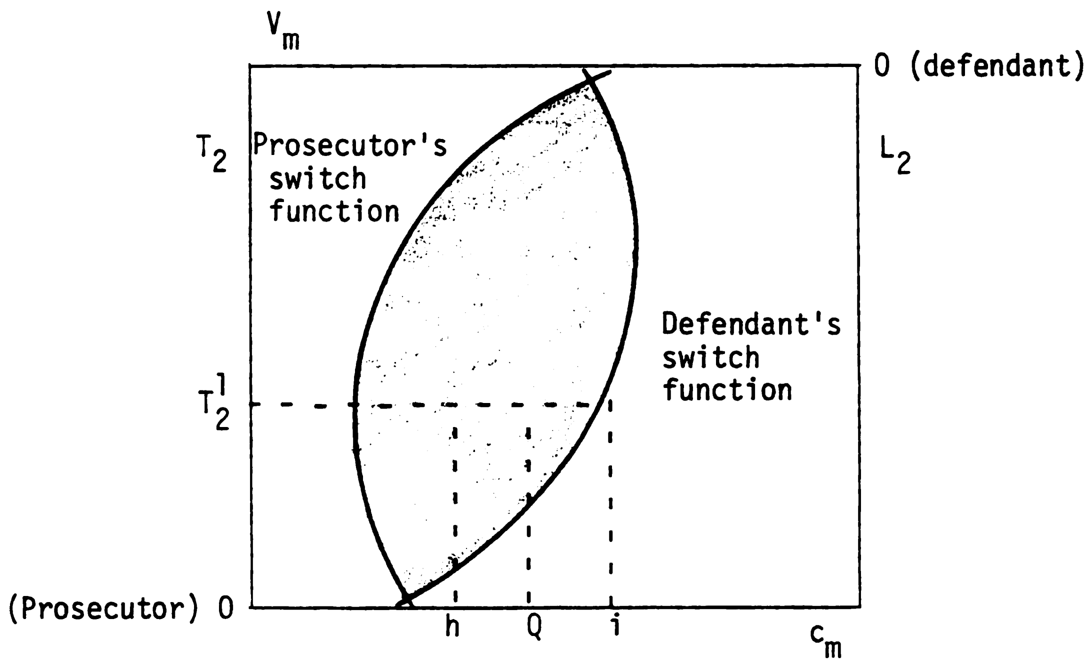



Figure 45.--Switch functions of prosecutor and defendant.

The simplifying assumption has been made that $T_2 = 1 - L_2$, i.e., that both parties agree that an increase in the strength of the other's case diminishes the strength of their own case and that what one case loses the other gains. The shaded area in Figure 45 is all those combinations of c and initial strength which can result in a bargain. It is assumed that if the prosecutor and defendant find themselves at points within the shaded area that they will reach a bargain.

## 4.14 The Probability of Bargaining

Since $0 \leq V_m \leq 1$, $0 \leq c_m \leq 1$, $0 \leq L_2^0 \leq 1$, $0 \leq T_2^0 \leq 1$, the area of the Edgeworth box is one. The probability of reaching a bargain is the area of the shaded region in Figure 45 divided by the area of the whole box or in this case simply the area of the shaded area. (The following discussion is from the vantage point of the prosecutor and thus all references to Figure 45 are made looking from the south-west corner of the box.) The probability of bargaining given $T_2^1 =$ $P(B/T_2^1) = i - h$. Given $T_2^1$ the case against the defendant will come to some conclusion and if all $c_m$ are considered equally likely then the probability of bargaining is simply the width of the shaded area at $T_2^1$, i.e., how much of the unit line segment represents bargains. This segment of the unit line segment is $i - h$. The probability that $c_{pd} > Q$ given bargaining and $T_2^1$ is $P((c_{pd} > Q)/B,T_2^1) = P((c_{pd} > Q),$ $B,T_2^1)/P(B,T_2^1) = \frac{i - Q}{i - h}$ if $Q < i$ otherwise zero. Thus, as expected, the higher Q is, the smaller the probability that the negotiated plea lies above it. (Conversely, the higher Q is, the greater the likelihood that the negotiated plea lies below it.)

## 4.15 The Bargaining Space and Shifts in the Level of Risk

Given dc/dr > 0, dV/dr > 0 it was shown in Chapters III and IV that both the prosecutor's and defendant's switch functions shift to the left, in the context of Figure 45, given the introduction of the mandatory sentence law.

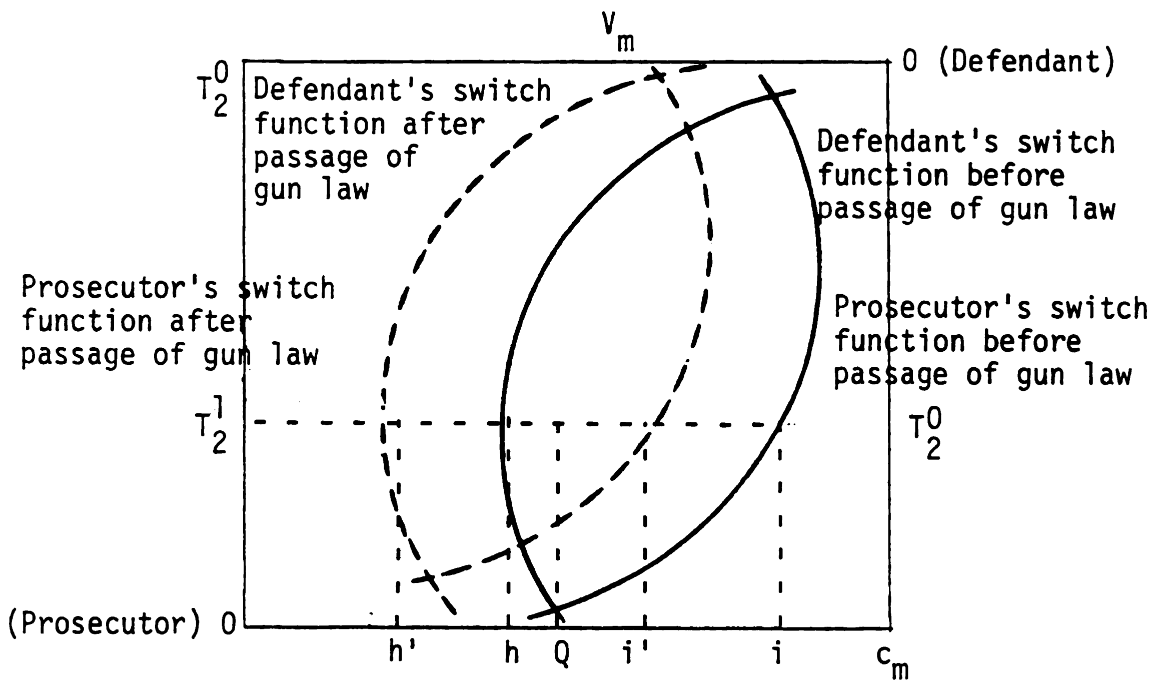



Figure 46.--Changes in the level of risk and shifts of the switch functions.

$P(c_{pd} > Q/B, \text{gun law}, T_2^1) = \frac{i' - Q}{i' - h'} < \frac{i - Q}{i - h}$ . This probability falls for two reasons. First, the maximum c that the defendant will plead guilty to (i') has fallen. This directly reduces the above probability. Second, the minimum c that the prosecutor will accept given $T_2^1$ has fallen. This reduction in the minimum c (h') increases the denominator of the above probability, causing it to fall. Thus passage of the gun law results in a reduction in the probability that the agreed-upon charge will be greater than Q. The implication is that there should be a general reduction in the severity of plea bargained charges since passage of the gun law.

## 4.16 The Other Cases

The change in the mandatory gun law can result in four possible changes in the bargaining space of the prosecutor and defendant, one of which, the one felt to be most likely, is depicted in Figure 46 of Section 4.15. The four cases can be categorized by the signing of dc/dr and dV/dr.

| | | |
|---|---|---|
| Case 1 | $dc/dr > 0$ | $dV/dr > 0$, |
| Case 2 | $dc/dr > 0$ | $dV/dr < 0$, |
| Case 3 | $dc/dr < 0$ | $dV/dr > 0$, |
| Case 4 | $dc/dr < 0$ | $dV/dr < 0$. |

### Case 1, $dc/dr > 0$, $dV/dr > 0$

Case 1 describes the conditions that are generally discussed within the dissertation. In this case it will be remembered the introduction of the gun law resulted in the bargaining space shifting to the left. Two results that can be drawn from Figure 47 are:

i) P(bargaining/gun law) $\gtrless$ P(bargaining/no gun law),

ii) P($c_{pd} \geq Q/T_2^0$, gun law) < P($c_{pd} \geq Q/T_2^0$, no gun law).

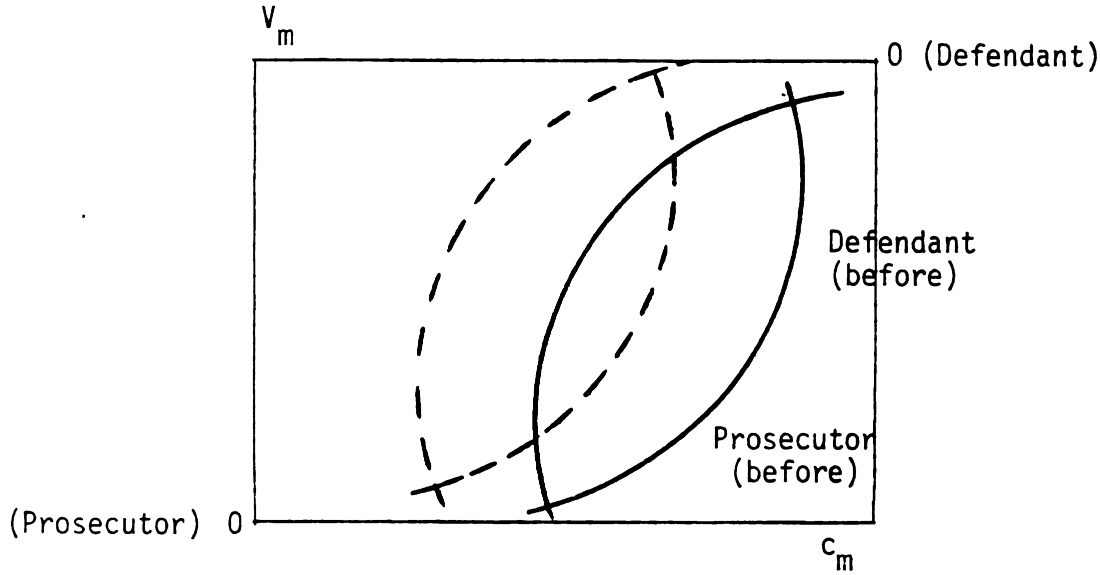



Figure 47.--Case 1.

Case 2, $dc/dr > 0$, $dV/dr < 0$

In case 2 the prosecutor is as before; the defendant, however, instead of resisting the increase in the level of risk to himself by attempting to acquire a higher $V_m$ is willing to settle for a lower $V_m$ as a result of the passage of the gun law.[16] $dV/dr$ will be unambiguously negative if $U_{VZZ} < 0$ for all Z and $U_{KZZ} > 0$ for all Z.[17] This would occur as V and Z are "complements" and K and Z are "substitutes."[18] In case 2 where $dc/dr > 0$ and $dV/dr < 0$ it follows that:

i) P(bargaining/gun law) > P(bargaining/no gun law),

ii) $\overline{P} = P(c_{pd} \geq Q/T_2^0,$ gun law$) \gtreqless P(c_{pd} \geq Q/T_2^0,$ no gun law$) = \underline{P}$.

Since the area between the dashed curves is greater than the area between the solid curves, the likelihood of bargaining also increases. It is impossible to a priori sign $(\overline{P} - \underline{P})$ since the defendant is now willing to accept a higher charge than prior to passage of the gun law while at the same time the prosecutor is willing to accept a lower charge. The sign of $(\overline{P} - \underline{P})$ depends upon which switch function shifts most because of the change in the level of risk.

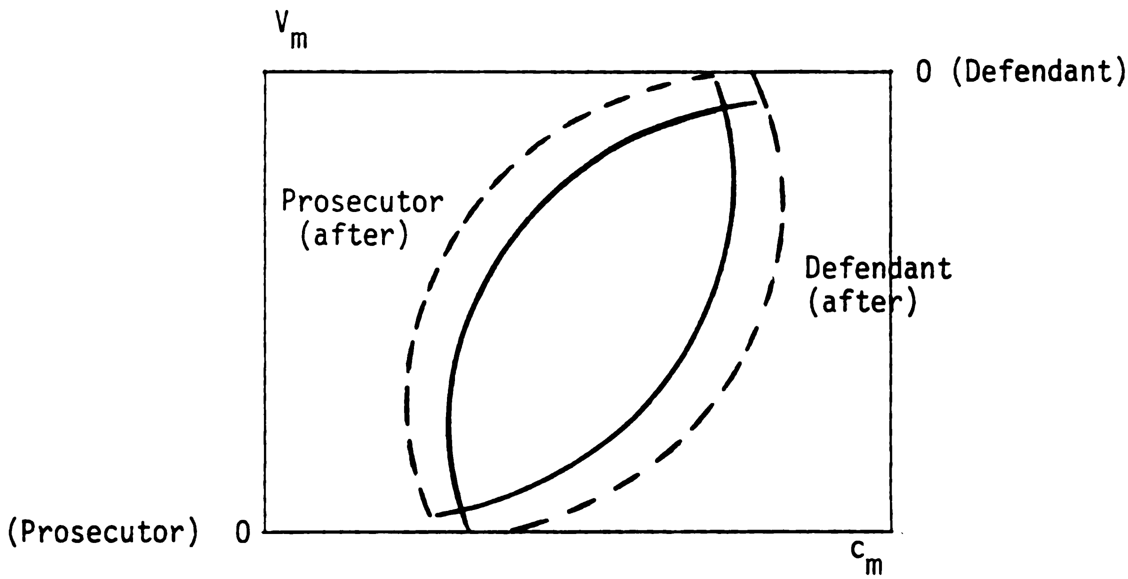



Figure 48.--Case 2.

Case 3, $dc/dr < 0$, $dV/dr > 0$

In this case $dV/dr$ takes the typically assumed sign. The prosecutor now responds to the reduction in the level of risk by seeking an increase in c. This would occur unambiguously if c and Y are viewed by the prosecutor as "complements" (rather than substitutes) and p and Y are viewed as "substitutes" (rather than complements).[19] In this case it follows that:

i) P(bargaining/gun law) < P(bargaining/no gun law),

ii) $\overline{P} = P(c_{pd} \geq Q/T_2^0, \text{ gun law}) \gtrless P(c_{pd} \geq Q/T_2^0, \text{ no gun law}) = \underline{P}$.

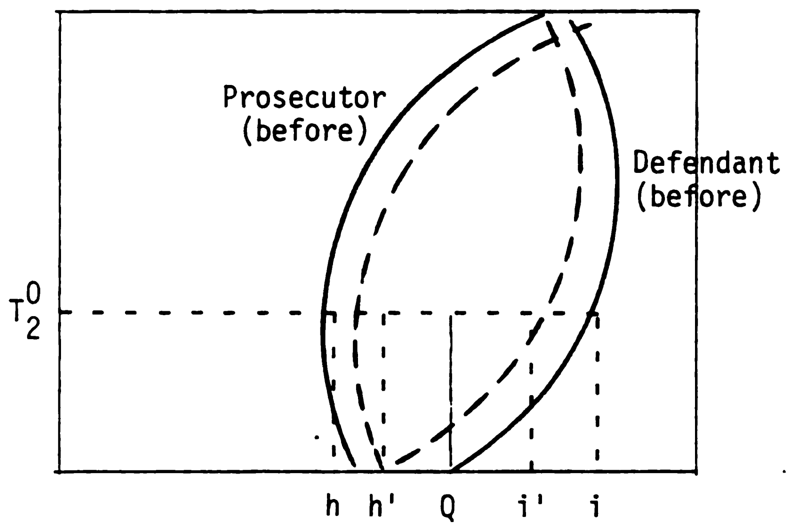



Figure 49.--Case 3.

The ambiguity of the sign of $(\overline{P} - \underline{P})$ can be shown as follows. The probability $P(c_{pd} > Q/T_2^0$, no gun law$) = \frac{i - Q}{i - h}$ as seen in Figure 49, while $P(c_{pd} > Q/T_2^0$, gun law$) = \frac{i' - Q}{k' - h'} = \frac{i - \Delta_i - Q}{i - \Delta_i - (h + \Delta_h)} =$

$$\frac{i - Q - \Delta_i}{(i - h) - \Delta_i - \Delta_h} = \frac{\frac{i - Q}{i - h} \quad \frac{\Delta_i}{i - h}}{1 - \frac{(\Delta_i + \Delta_h)}{i - h}} = \frac{\underline{P} - \frac{\Delta_i}{i - h}}{\frac{1 - (\Delta_i + \Delta_h)}{i - h}}$$

$\Delta_i = i - i' > 0$ and $\Delta_h = h' - h > 0$.

Now $\text{sgn}(d\overline{P}) = d\Delta_i \left\{ 1 - \frac{(\Delta_i + \Delta_h)}{i - h} \right\} + (d\Delta_i + d\Delta_h) \left\{ \underline{P} - \frac{\Delta_i}{i - h} \right\}$.

It follows that $d\overline{P} < 0$ when $\overline{P} \quad \frac{d\Delta_i}{d\Delta_i + d\Delta_h}$ . .

Note that if the defendant is insensitive to the change in the level of risk and $dc/dr < 0$ then the condition is never met and $\underline{P} < \overline{P}$.

Conversely, if the prosecutor is insensitive to changes in the level of risk, i.e., $d\Delta_h = 0$ and $dV/dr > 0$, then $d\overline{P}$ is always negative, i.e., $P(c_{pd} \geq Q/T_2^0$, gun law) < $P(c_{pd} \geq Q/T_2^0$, no gun law).

Case 4, $dc/dr < 0$, $dV/dr < 0$

This case is the exact opposite of case 1 in that both switch functions shift to the right rather than the left because of the change in the level of risk.

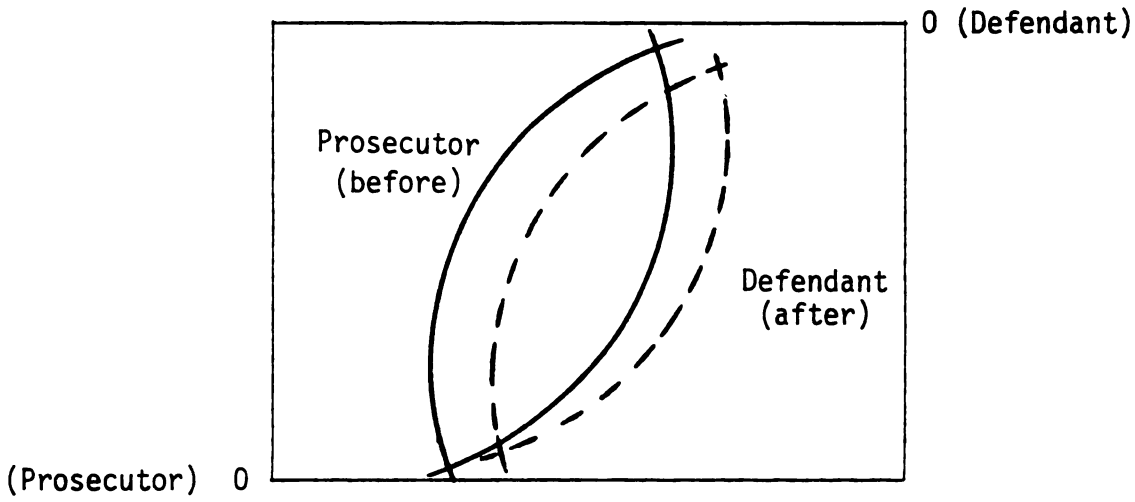



Figure 50.--Case 4.

In this case it follows that:

i) P(bargaining/no gun law) $\gtrless$ P(bargaining/gun law),

ii) $P(c_{pd} \geq Q/T_2^0$, no gun law) < $P(c_{pd} \geq Q/T_2^0$, gun law).

## 4.17 Separation of Cases

To uniquely determine which of the four cases above holds, it is necessary to determine three conditions (this is assuming that dc/dr and dV/dr are not both directly determined). They are:

1. The direction of change in the probability of bargaining;
2. Whether the probability that $c_{pd} > Q$, given $T_2^0$, has risen or fallen given the change in the level of risk;
3. The sign of either dc/dr or dV/dr. This evaluation must be performed only if $P(B/GL) \neq P(B/NGL)$.

These are sufficient since each condition excludes at least one of the four cases from being relevant. Thus a situation is never reached where it is impossible to distinguish between two or more cases because of a lack of uniquely identifying characteristics.

## 4.18 Example of Defendant's Maximization Problem

Assume that the defendant's behavior is represented by the following utility function:

$$U = \int_{\underline{Z}}^{\overline{Z}} \ln(1 + V) Z^{\alpha} G_Z + \int_{\underline{Z}}^{\overline{Z}} \ln(1 + K) Z^{\varepsilon} G_Z \quad \text{then}$$

$$du/dV = \int_{\underline{Z}}^{\overline{Z}} \frac{1}{(1 + V)} Z^{\alpha} G_Z > 0$$

$$d^2U/dV^2 = \frac{\int_{\underline{Z}}^{\overline{Z}} -Z^{\alpha} G_Z}{(1 + V)^2} < 0$$

$$du/dK = \int_{\underline{Z}}^{\overline{Z}} \frac{1}{(1 + K)} Z^{\varepsilon} G_Z > 0$$

$$d^2U/dK^2 = \frac{\int_{\underline{Z}}^{\overline{Z}} -Z^{\varepsilon}G_Z}{(1 + K)^2} < 0.$$

$$\frac{d^2U}{dVdZ} = \frac{\int_{\underline{Z}}^{\overline{Z}} \alpha Z^{\alpha-1}G_Z}{(1 + V)} < 0 \Rightarrow \alpha < 0$$

$$\frac{d^3U}{dVdZ^2} = \frac{\int_{\underline{Z}}^{\overline{Z}} \alpha(\alpha - 1)Z^{\alpha - 2}G_Z}{(1 + V)} > 0$$

$$\frac{d^2U}{dKdZ} \qquad \frac{\int_{\underline{Z}}^{\overline{Z}} \varepsilon Z^{\varepsilon - 1}G_Z}{(1 + K)} > 0 \Rightarrow \varepsilon > 0$$

$$\frac{d^3U}{dKdZ^2} = \frac{\int_{\underline{Z}}^{\overline{Z}} \varepsilon(\varepsilon - 1)Z^{\varepsilon-2}G_Z}{(1 + K)} < 0 \Rightarrow \varepsilon < 1 \qquad \text{thus } 0 < \varepsilon < 1.$$

Let $\alpha = -.4$, $\varepsilon = .1$ then V and Z are "substitutes" and K and Z are "complements." Assume that the defendant has a total budget (M) of \$100 with the average price of a unit of $V(P_V)$ being \$50 and the average price of a unit of consumption $(P_K)$ being \$1. The linear approximation to the defendant's budget constraint is

$$100 = 50V + 1K \quad \text{or} \quad M = P_V V + P_K K.$$

Assume that the defendant knows that V will be between $\bar{\bar{Y}} - b$ and $\bar{\bar{Y}} - a$ and believes that all values between these two extremes are equally likely. The uniform distribution is in this case an adequate representation of the random variable Z. Under this assumption then

$$F(L) = \int_{\underline{Z}}^{1} \frac{1\, dZ}{\overline{Z} - \underline{Z}} = \frac{1 - \underline{Z}}{\overline{Z} - \underline{Z}} = \frac{1 - \underline{Z}}{b - a} = P(Z < 1), \quad \underline{Z} \le 1 \le \overline{Z}$$

and $F_Z = 1/(\bar{z} - \underline{z}) = 1/(b - a)$.

In keeping with prosecutors' maximization problem let $\bar{\bar{Y}} = b = 5$ and $a = 1$. Thus the initial maximization problem is

$$\max_{V,K} \quad \int_{\underline{z}}^{\bar{z}} \frac{\ln(1 + V)Z^{-.4}}{(5 - 1)} + \int_{\underline{z}}^{\bar{z}} \frac{\ln(1 + K)Z^{.1}}{(5 - 1)} \quad \text{subject to}$$

$100 - 50V - 1K = 0$. Or it may be written as

$$\int_{\underline{z}}^{\bar{z}} \frac{\ln(1 + V)Z^{-.4}}{4} + \int_{\underline{z}}^{\bar{z}} \frac{\ln(1 + K)Z^{.1}}{4} - \lambda(100 - 50V - 1K).$$

The first order conditions will be written and solved using the symbolic values $\alpha$, $\varepsilon$, $P_V$, $P_K$, M in order to facilitate the determination of alternate values of V and K given different prices and levels of risk. The first order conditions for a maximum are:

$$\int_{\underline{z}}^{\bar{z}} \frac{Z^{\alpha}}{(1 + V)} + \lambda P_V = 0\,, \tag{49}$$

$$\int_{\underline{z}}^{\bar{z}} \frac{Z^{\varepsilon}}{(1 + K)} + \lambda P_K = 0\,, \tag{50}$$

$$M - P_V V - P_K K = 0. \tag{51}$$

Solving 50 for $\lambda$ and substituting into 49

$$\int_{\underline{z}}^{\bar{z}} \frac{Z^{\alpha}}{(1 + V)} - \frac{P_V}{P_K} \int_{\underline{z}}^{\bar{z}} \frac{Z^{\varepsilon}}{(1 + K)} = 0. \tag{52}$$

Solving 51 for K, substituting into 4.12.4 and performing the indicated integration results in

$$\frac{1}{(1 + V)} \frac{(\bar{z}^{\alpha+1} - \underline{z}^{\alpha+1})}{(1 + \alpha)} - \frac{P_V(\bar{z}^{\varepsilon+1} - \underline{z}^{\varepsilon+1})}{P_K(1 + K)(1 + \varepsilon)} = 0. \tag{53}$$

Solving for V results in

$$V = \frac{(\varepsilon + 1)(M + 1)(\overline{z}^{\alpha+1} - \underline{z}^{\alpha+1}) - P_V(\overline{z}^{\varepsilon+1} - \underline{z}^{\varepsilon+1})(\alpha + 1)}{P_V((\varepsilon + 1)(\overline{z}^{\alpha+1} - \underline{z}^{\alpha+1}) + (\alpha + 1)(\overline{z}^{\varepsilon+1} - \underline{z}^{\varepsilon+1}))} . \tag{54}$$

In our example then where $\alpha = -.4$, $\varepsilon = .1$, $P_V = \$50$, $P_K = \$1$, $M = \$100$, $a = 1$, $b = 5$ and $\bar{\bar{Y}} = 5$ then

$$V^* = \frac{(1.1)(101)(4^{.6} - 0) - 50(4^{1.1})(.6)}{50((1.1)(4^{.6}) + (.6)(4^{1.1}))}$$

$$V^* = \frac{2.55.24 - 137.84}{50(2.52 + 2.76)} = .44$$

It follows that the optimal K, K* = 77.78. The level of total utility given the optimal V and K is 4.91.

Given the mandatory sentence law (derived by letting a = 3), the optimal solution is

$$V^* = \frac{(1.1)(101)(2^{.6}) - 50(2^{1.1})(.6)}{50((1.1)(2^{.6}) + .6(2^{1.1}))} = .70 .$$

It follows that the optimal K, K* is 64.76 and total utility given the optimal V and K is 4.75.

The above solutions were developed under the assumption that resolution would be via trial. The following two solutions are developed given bargaining in the pre and post gun law cases. In the pre-gun law case assume that bargaining raises the price of an average unit of V from \$50 to \$75. All other variables remain at the levels in the pre-gun law example above. Then

$$V^* = \frac{(1.1)(101)(4^{.6}) - 75(4^{1.1})(.6)}{75((1.1)(4^{.6}) + (.6)(4^{1.1}))} = .12$$

It follows that the optimal K, $K^* = 90.77$ and that total utility given the optimal V and K is 4.83.

If the change in price were the only effect of bargaining, the defendant in this case would not bargain since total utility is higher in the trial case (4.91 vs. 4.83). It was argued in the main body of the text that bargaining also results in an increase in the maximum level of consumption that the defendant can obtain. In the model this effect is represented by an increase in M. If bargaining results in the effective M increasing by at least 8 then the defendant will find bargaining desirable. Letting M increase from 100 to 108 and keeping $P_V = 75$, ceteris paribus results in

$$V^* = \frac{(1.1)1(109)(4^{.6}) - 75(4^{1.1})(.6)}{75((1.1)(4^{.6}) + (.6)(4^{1.1}))} = .17$$

In this case the optimal K, $K^*$ is 94.93 and the total level of utility given the optimal level of K and V (and the new level of M) is 4.91. This is the same level of utility as in the trial version of the problem. In this particular case the defendant is indifferent between trial and bargaining. If the price of V were to fall below 75 or the increase in M were greater than \$8, then the defendant would prefer to bargain rather than to go to trial.

In the post-gun law example it is again assumed that bargaining raises the average price of V to \$75 (in this case a = 3). All other variables take on the values they had in the post-gun law trial problem. The optimal V then is

$$V^* = \frac{(1.1)(101)(2^{.6}) - 75(2^{1.1})(.6)}{75((1.1)(2^{.6}) + (.6)(2^{1.1}))} = .43$$

In this case it follows that the optimal K, K* is 85.54 and the level of total utility given K* and V* is 4.57. If the change in the average price of $P_V$ were the only change then the defendant in this case would not bargain since the total utility in the trial model (given the gun law) exceeds the total utility of bargaining (4.75 vs. 4.57). In this case the increase in M necessary to make the defendant indifferent between bargaining and trial is $14. If M = $114 instead of $100 then

$$V^* = \frac{(1.1)(115)(2^{.6}) - 75(2^{1.1})(.6)}{75((1.1)(2^{.6}) + (.6)(2^{1.1}))} = .43$$

In this case the optimal K, K* is 81.74 and the level of utility given K* and V* is 4.75. If the price of V were to fall below $75 or the increase in M to exceed $14, then the defendant would prefer bargaining to trial.

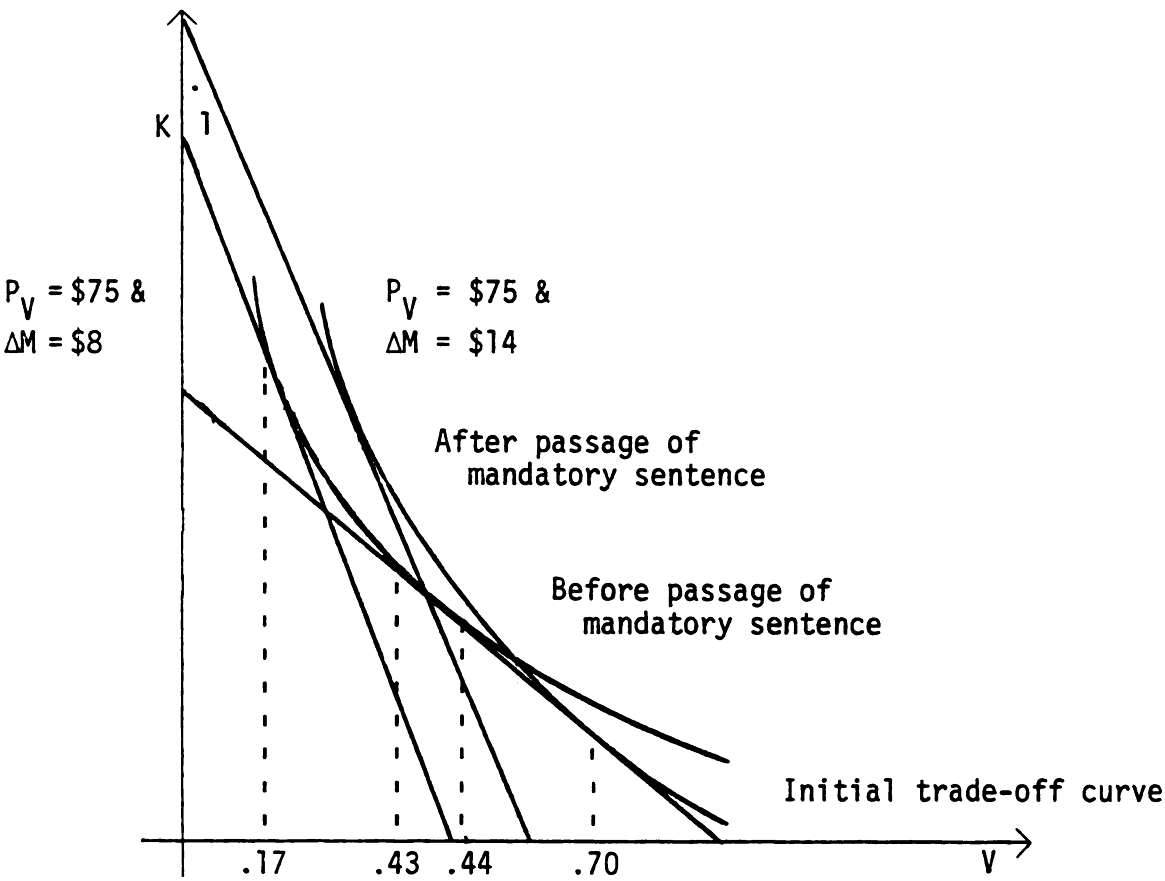



Figure 51.--The various cases.

## Footnotes--Chapter IV

[1]Wealth in fact is not a constant since from time of arrest to trial can be a considerable time. The amount of wealth available for preparation of a defense depends heavily upon the availability of bail. Appropriately, if wealth effects were to be a major area of concern W should be modeled as:

$$W = W_1 - C_b + \int_0^T E(Z)dZ \quad \text{where}$$

$W_0$ is the wealth level at time of arrest; $C_b$ is the cost of bail; $E(Z)$ is a daily earnings function and $(0,T)$ is the total time from arrest to conviction. See Richard P. Adelstein, "The Plea Bargain in Theory: A Behavioral Model of the Negotiated Guilty Plea," The Southern Economic Journal 6 (January 1978): 495. See also William M. Landes, "An Economic Analysis of the Courts," The Journal of Law and Economics 14 (April 1971): 61-108.

[2]$\bar{\bar{Y}}$ is the estimated maximum sentence that the defendant believes he can receive. As in the chapter on the prosecutor, a is the minimum sentence and b is the maximum sentence that the defendant can receive. It is assumed, for simplicity, that $\bar{\bar{Y}} - b$ is zero. Z is a measure of how far the actual sentence is from the maximum sentence. It is assumed that larger values of Z are preferred to smaller values.

[3]I.e., V and K are not stochastic while Z is.

[4]The most current example appears in the State Journal (Lansing, Michigan), January 5, 1979, sec. 1. In this account the judge refused to accept an agreed-upon (by the prosecutor and defendant) maximum sentence of 10 years.

[5]This is not exact since the original definition of risk aversion is for a utility function with one commodity. For an extension of the definition of a utility function with several commodities, see R. Kihlstrom and Leonard Mirman, "Risk Aversion with Many Commodities," Journal of Economic Theory 8,3 (1974): 361-88. In this particular case the defendant is risk averse if

$$\frac{\begin{vmatrix} U_{KK} & U_{KV} \\ U_{VK} & U_{VV} \end{vmatrix}}{\begin{vmatrix} U_{KK} & U_{KV} & U_K \\ U_{KV} & U_{VV} & U_V \\ U_K & U_V & 0 \end{vmatrix}^{\frac{2}{3}}} > 0$$

[6]This classification is correct in the Edgeworth-Pareto sense. For an application of this approach see R. G. Lipsey and Kelvin Lancaster, "The General Theory of the Second Best," Review of Economic Studies 24 (July 1956-57): 11-32, esp. p. 29.

[7]Kelvin Lancaster, Consumer Demand, A New Approach (New York: Columbia University Press, 1971).

[8]r is the level of risk and is explained in Appendix A.11 and A.12. The derivation of dV/dr is in Section 4.6.

[9] For an explanation of the Lagrange Multiplier see Alpha C. Chiang, Fundamental Methods of Mathematical Economics (New York: McGraw-Hill, 1967), pp. 350-54.

[10]Taro Yamane, Mathematics for Economists, 2nd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1968), pp. 475-80.

[11] $\int_{\underline{z}}^{\bar{z}} U_V G_{Zr} = U_V G_V \Big|_{\underline{z}}^{\bar{z}} - \int_{\underline{z}}^{\bar{z}} U_{VZ} G_r = \int_{\underline{z}}^{\bar{z}} U_{VZZ} \int_0^y G_r \;\forall\; y \in [\underline{z}, \bar{z}]$.

From Appendix A.11 and A.12 $\int_0^y G_r \geq 0 \;\forall\; y \in [\underline{z}, \bar{z}]$, thus $\mathrm{sgn} \int_z^z U_V G_{Zr} = \mathrm{sgn} U_{VZZ}$.

[12]For a discussion of transformations of variables and that Z is closed under convolution see D. A. S. Fraser, Probability and Statistics (North Scituate: Duxbury Press, 1976), p. 115.

[13]This is a direct application of Stochastic Dominance. See Appendix A.1 through A.7.

[14]Josef Hadar, Mathematical Theory of Economic Behavior (Reading, Mass.: Addison-Wesley, 1971), p. 179.

[15]Ibid., p. 191.

[16]This could be viewed as the "handwriting on the wall" case.

[17]See Section 4.5.

[18]In general dV/dr < 0 if dMRS/dr < 0 (see Section 4.7). The Marginal Rate of Substitution =

$$\int_z^z U_V G_Z / \int_z^z U_K G_Z = -dK/dV$$

and is a measure of how K and V can be substituted for each other while leaving the defendant at the same level of Utility. The MRS falling states that a small increase in V must cost less in terms of consumption than before the increase in the level of risk, if it is to leave the defendant at the same level of utility. This could happen if the defendant saw his case as "hopeless" since his is to receive a two-year mandatory sentence over a wide range of V. Thus he is less willing to give up consumption in return for an increase in V.

[19] In general for $dc/dr < 0$ it is required that $dMRS/dr < 0$ where MRS is the marginal rate of substitution = $dp/dc$. This could occur if the reduction in the level of risk increased the utility of c (at the margin relative to p). Thus a small reduction in c now requires a larger compensating increase in p (relative to before the change in the law) if the prosecutor is to maintain the same level of utility.

## CHAPTER V

## DATA EVALUATION

### 5.1 Purpose and Background

On January 1, 1977, Michigan House Bill No. 5073 became effective. It mandates at least two years in prison for any person convicted of a felony while in possession of a firearm. House Bill No. 5073 is also referred to as the gun law. Although the law addresses particular types of crime, it may also be considered a limited experiment in mandatory sentencing. The purpose of this study is to provide an evaluation of the gun law.

### 5.2 Gun Law Study Data Source

All probation officers in the state of Michigan were required to send the basic information sheet, the case summary, and the investigator's version and/or police officer's arrest report for persons sentenced on a felony during the sample time periods.[1] The time periods consisted of one two-week period in 1976 (before the gun law was in effect) and one two-week time period in late 1977 (after the gun law). The information obtained from these sheets includes original charge, the charge on which conviction was obtained, min/max sentence, sentencing judge, retained or appointed lawyer, age, race, sex, level of education, employment history, previous criminal record, marital status, number of children, and history of drug, alcohol, or sex abuses. Also obtained was a short description of the crime.

## 5.3 Sample Characteristics

Tables 1 and 2 compare the samples on selected crime and sentencing variables. The samples are similar on most variables. Within offense type, small differences occur in robbery, property, and drug offenses. Robbery and drug offenses, as a percentage of total persons sentenced in the time periods, were higher in 1976. On the other hand, property crimes increased in 1977, as a proportion of total offenses committed. As Table 1 indicates, firearms were involved in more offenses during 1977 than 1976. The major differences in dispositions are that more people received jail sentences in 1977 and less probation. For the majority of both samples, the length of sentence is almost identical. The minimum prison sentence in both 1977 and 1976 was approximately 2-1/2 years.

## 5.4 General Findings

Since the gun law excludes those persons convicted of carrying a concealed weapon (CCW, Michigan Criminal Law No. 227), they are included in the non-firearm offense category.

Table 3 compares the possession of firearms during the commission of a non-CCW felony for the 1976 and 1977 samples. The results indicate no significant differences between time periods in the use of a firearm. Finding no difference suggests that the gun law does not have an impact on the use of a firearm during the commission of a non-CCW felony.

Table 1.--Sample characteristics: frequency and percentage distribution of crime data for 1976 and 1977.

| Item | 1976 | | 1977 | |
|---|---|---|---|---|
| | N (683) | % | N (540) | % |
| Offense Type | | | | |
| Homicide | 20 | 2.9 | 11 | 2.0 |
| Robbery | 70 | 10.3 | 36 | 6.7 |
| CSC | 15 | 2.2 | 18 | 3.3 |
| Assault | 36 | 5.3 | 28 | 5.2 |
| Children & sex | 7 | 1.0 | 0 | 0.0 |
| Property | 383 | 56.0 | 341 | 63.0 |
| Drugs | 88 | 12.9 | 54 | 10.0 |
| Weapons | 53 | 7.8 | 43 | 8.0 |
| Other | 9 | 1.3 | 9 | 1.7 |
| Total | 681 | 100.0 | 540 | 100.0 |
| Weapon | | | | |
| None | 512 | 75.2 | 399 | 74.0 |
| Firearm* | 117 | 17.2 | 106 | 19.7 |
| Knife | 30 | 4.4 | 18 | 3.3 |
| Other | 22 | 3.2 | 16 | 3.0 |
| Total | 681 | 100.0 | 539 | 100.0 |

*Does not include accomplices.

Table 2.--Sample characteristics: frequency and length of sentence for 1976 and 1977.

| Disposition | 1976 | 1977 |
|---|---|---|
| Prison Min/Max[2] | | |
| Number | 229 | 179 |
| % of total | 33.5 | 33.1 |
| Median | 2 yr. 6 m.-7 yr. | 2 yr. 6 m.-5 yr. 2 m. |
| Probation | | |
| Number | 416 | 315 |
| % of total | 60.9 | 58.3 |
| Median | 24 m. | 25 m. |
| Jail | | |
| Number | 33 | 42 |
| % of total | 4.8 | 7.8 |
| Median | 11 yr. 5-1/2 m. | 11 yr. 5 m. |
| Other | | |
| Number | 5 | 4 |
| % of total | 0.7 | 0.7 |

Table 3.--Comparison of 1976 and 1977, the possession of firearms in non-CCW felonies as a percentage of all persons sentenced.

| | 1976 | | 1977 | |
|---|---|---|---|---|
| | N | % | N | % |
| Non-firearm offenses | 595 | 87.4 | 472 | 87.6 |
| Firearm offenses | 86 | 12.6 | 67 | 12.4 |
| Total | 681 | 100.0 | 539 | 100.0 |

Although the gun law does not appear to have an impact on the use of firearms, a related issue concerns the law's possible effect on non-firearm weapons. It appears that a possible outcome of the gun law would be an increase in other types of weapons used during commission of felonies. The samples do not support such a hypothesis. Tables 4 and 5 below clarify this point.

Table 4.--The use of non-firearm weapons during the commission of a non-CCW offense by year.

| | 1976 | | 1977 | |
|---|---|---|---|---|
| | N | % | N | % |
| Non-firearm weapon offenses | 44 | 6.5 | 32 | 5.9 |
| All other offenses | 637 | 93.5 | 507 | 94.1 |
| Total | 681 | 100.0 | 539 | 100.0 |

Table 5.--Summary of weapon involvement and firearm felony convictions for 1976 and 1977.

| | 1976 | | 1977 | |
|---|---|---|---|---|
| | N | % | N | % |
| Total number sentenced | 681 | | 539 | |
| Involved in a firearm offense* | 86 | 12.6 | 67 | 12.4 |
| Had other weapon in possession* | 44 | 6.5 | 32 | 5.9 |
| % with firearm in possession who were convicted under the gun law | n/a | | 19 | 28.4 |

*These categories exclude CCW offenses only.

The contention of this dissertation is that passage of the gun law would result in reduced sentences for those convicted of using a gun during the commission of a felony. If the charge is reduced in those cases involving a gun, then the minimum sentence should fall. This is borne out by Table 6.

Table 6.--Comparison of median prison sentence for persons involved in a firearm offense (non-CCW) in 1976 and 1977 with those persons convicted of the gun law in 1977 (not including the two-year add-on).

| Prison | 1976 | 1977 | |
|---|---|---|---|
| (Minimum Sentence) | Firearm Involvement | Firearm Involvement No Gun Law Conviction | Convicted of the Gun Law |
| Number | 59 | | 19 |
| Median | 5 yrs., 1 m. | 5 yrs., 0 m. | 3 yrs., 2 m. |

The results show the median prison sentence given for those persons convicted of the gun law is almost two years less than those convicted of a firearm offense in 1976.[3] With the two-year add-on considered (all persons sentenced on the gun law in this sample received the two-year add-on), the only extension in time served is primarily a function of good-time loss.[4] This finding supports the conclusion that the gun law does not have an impact on the length of sentence. The most interesting observation in Table 7 is the longer median prison sentence received for those not convicted of the gun law but involved in violent firearm offenses during 1977.

Table 7.--Comparison of median prison sentences for non-CCW firearm involvement in violent and non-violent offenses in 1976 and 1977.

| Prison (Minimum Sentence) | 1976 | 1977 | |
|---|---|---|---|
| | Firearm Involvement | Firearm Involvement No Gun Law Conviction | Convicted of the Gun Law |
| Violent | | | |
| Number | 55 | 13 | 19 |
| Median | 5 yrs., 1 m. | 6 yrs., 6 m. | 3 yrs., 2 m. |
| Non-violent | | | |
| Number | 4 | 8 | 0 |
| Median | 3 yrs., 2 m. | 3 yrs., 0 m. | - |

A final issue concerning the general impact of the gun law concerns sentencing patterns. The type of sentence received for offenses involving a weapon is shown in Table 8. Of those persons sentenced for crimes involving firearms during 1977, 59.7% were given prison terms. By comparison, 69.4% of those persons sentenced during 1976 for firearm offenses received prison. Therefore, fewer people involved in firearm crimes received prison during 1977 than in 1976. This is consistent with the results of Table 6, which indicated the general reduction in sentences during 1977 relative to 1976 for crimes involving a gun. Table 7 indicates that increased use of probation occurred also for crimes that included the use of a weapon other than a gun.

Table 8.--Dispositions by weapon and year.

| | 1976 | | | | | | 1977 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | None | | Firearms | | Other | | None | | Firearms | | Other | |
| | N | % | N | % | N | % | N | % | N | % | N | % |
| Prison | 144 | 26.1 | 59 | 69.4 | 26 | 59.1 | 126 | 28.6 | 40 | 59.7 | 13 | 40.6 |
| Probation | 378 | 68.5 | 20 | 23.6 | 17 | 38.6 | 276 | 62.8 | 22 | 32.9 | 17 | 53.1 |
| Other | 30 | 5.4 | 6 | 7.0 | 1 | 2.3 | 38 | 8.6 | 5 | 7.4 | 2 | 6.9 |
| Total | 552 | | 85 | | 44 | | 440 | | 67 | | 32 | |

## 5.5 Summary of Gun Law Findings

1. There is no observable impact on the proportion of felonies in which a weapon is used, or the type of weapon (Tables 3, 4, and 5).

2. The results indicate the median sentence given for those persons convicted of the gun law is almost two years less than those convicted of a firearm offense in 1976. (The two-year add-on is not calculated in the above figures.) With the two-year add-on considered, the impact of the length of sentence is negligible (Table 6).

3. The above conclusion continues to be supported when the offenses are aggregated into violent/non-violent categories (Table 7).

4. A related finding is the longer median sentence received for those persons convicted of a firearm offense during 1977 but not receiving the gun law add-on (Table 7).

5. Dispositions given for firearm offenses during 1977 and 1976 have been examined. Of those persons sentenced for crimes involving firearms during 1977, 59.7% were given prison terms. By comparison, 69.5% of those persons sentenced during 1976 for firearm offenses received prison terms (gun law convictions are included in these figures). Fewer people involved in firearm crimes received prison terms during 1977 than in 1976. The same observation can be made with respect to other types of weapons (Table 8) but not to sentences in general (Table 2).

## Footnotes--Chapter V

[1]These data were sent to the Michigan Department of Corrections, Program Bureau, Lansing, Michigan. The Program Bureau is responsible for the collection and tabulation of data in this chapter.

[2]Time chart for minimum prison terms, Michigan Department of Corrections, Program Bureau, March 2, 1977:

TIME TO BE SERVED

| Minimum Terms | | With Good Time Allowances | | | | | |
|---|---|---|---|---|---|---|---|
| | | Regular | | | Special | | |
| Yrs. | Mos. | Yrs. | Mos. | Days | Yrs. | Mos. | Days |
| - | 6 | - | 5 | - | - | 4 | 15 |
| 1 | - | - | 10 | 0 | 0 | 9 | - |
| 1 | 6 | 1 | 3 | - | 1 | 1 | 15 |
| 2 | - | 1 | 8 | - | 1 | 6 | - |
| 2 | 6 | 2 | - | 24 | 1 | 10 | 6 |
| 3 | - | 2 | 5 | 18 | 2 | 2 | 12 |
| 3 | 6 | 2 | 10 | 12 | 2 | 6 | 18 |
| 4 | - | 3 | 3 | 6 | 2 | 10 | 24 |
| 4 | 6 | 3 | 7 | 24 | 3 | 2 | 21 |
| 5 | - | 4 | - | 12 | 3 | 6 | 18 |
| 5 | 6 | 4 | 5 | - | 3 | 10 | 15 |
| 6 | - | 4 | 9 | 18 | 4 | 2 | 12 |
| 6 | 6 | 5 | 1 | 24 | 4 | 5 | 21 |
| 7 | - | 5 | 6 | - | 4 | 9 | - |
| 7 | 6 | 5 | 10 | 6 | 5 | - | 9 |
| 8 | - | 6 | 2 | 12 | 5 | 3 | 18 |
| 8 | 6 | 6 | 6 | 18 | 5 | 6 | 27 |
| 9 | - | 6 | 10 | 24 | 5 | 10 | 6 |
| 9 | 6 | 7 | 2 | 24 | 6 | 1 | 6 |
| 10 | - | 7 | 6 | 24 | 6 | 4 | 6 |
| 11 | - | 8 | 2 | 24 | 6 | 10 | 6 |
| 11 | 6 | 8 | 6 | 24 | 7 | 1 | 6 |
| 12 | - | 8 | 10 | 24 | 7 | 4 | 6 |
| 12 | 6 | 9 | 2 | 24 | 7 | 7 | 6 |
| 13 | - | 9 | 6 | 24 | 7 | 10 | 6 |
| 14 | - | 10 | 2 | 24 | 8 | 4 | 6 |
| 15 | - | 10 | 10 | 0 | 8 | 9 | - |
| 16 | - | 11 | 5 | 6 | 9 | 1 | 24 |
| 17 | - | 12 | - | 12 | 9 | 6 | 18 |
| 18 | - | 12 | 7 | 18 | 9 | 11 | 12 |
| 19 | - | 13 | 2 | 24 | 10 | 4 | 6 |
| 20 | - | 13 | 8 | 24 | 10 | 7 | 6 |
| 22 | 6 | 14 | 11 | 24 | 11 | 2 | 21 |
| 25 | - | 16 | 2 | 24 | 11 | 10 | 6 |
| 30 | - | 18 | 8 | 24 | 13 | 1 | 6 |
| 35 | - | 21 | 2 | 24 | 14 | 4 | 6 |

The time to be served with regular or special good time allowance is the absolute minimum time before eligibility for release on parole.

[3]It is readily admitted that in the crowded prison conditions that exist today, the length of time actually spent in jail is a function of the minimum sentence rather than the maximum sentence received. Toward this end the Michigan Department of Corrections has developed a "good-time" sheet used, among other things, to determine probable release dates of convicts. This chart is prepared using information on minimum sentence only; that is, maximum sentence is not a factor in determining release date.

[4]The two-year mandatory sentence imposed by the gun law must be served prior to serving time for other convictions; that is, gun time is not served consecutively with other time. Using the time and cost chart below, it can be calculated that a person receiving a 5-year minimum should expect to serve approximately 4 years and 12 days in jail. If convicted of the gun law the minimum sentence implies a jail term of 2 years and 5 months. To this must be added the two-year mandatory sentence for using a gun. This results in a total prison time of 4 years and 5 months. If the convict receives good time, this sentence is reduced to 4 years and 2 months. This would tend to support the contention that the mean of Y has remained the same if only slightly increased.

## CHAPTER VI

## EMPIRICAL TESTING AND RESULTS

The empirical/statistical verification of the model requires up to three tests in order to uniquely identify the behavior of the prosecutor and defendant to changes in the level of risk.[1] In this chapter the tests to be performed are outlined, the econometric specifications given and results interpreted.

### 6.1 Test for Effect of Change in the Level of Risk on the Probability of Bargaining

It was shown in Chapter IV that three tests, at most, are necessary to discriminate between the four cases as outlined in Section 4.16. The four cases can be summarized as follows:

Case 1: $dc/dr > 0$, $dV/dr > 0$,
Case 2: $dc/dr > 0$, $dV/dr < 0$,
Case 3: $dc/dr < 0$, $dV/dr > 0$,
Case 4: $dc/dr < 0$, $dV/dr < 0$.

The first test to be performed is to estimate the effect of the gun law on the probability of bargaining. In this test

$$P_B = P(\text{Bargaining}) = f(c,cG,G,A,E,R,L,M,c^2,c^2G,\ D,ED,RE,\ DEL,P,DRG,S,MEN,AL) \tag{55}$$

c is the Sellin-Wolfgang measure of crime severity.[2]

G is equal to one if the crime was committed before passage of the mandatory gun law and zero otherwise.

A is one if the defendant is 17 or younger and zero otherwise.

E is one if the defendant was employed at the time of arrest and zero otherwise.

R is one if the defendant is non-Caucasian and zero otherwise.

L is one if defense council is appointed and zero otherwise.

M is one if the defendant is married and zero otherwise.

D is equal to one if the defendant was tried within the Detroit SMSA and zero otherwise

ED is equal to the number of formal years of education completed by the defendant

RE is equal to one if the victim is an acquaintance of the defendant and zero otherwise

DEL is the number of days between the date of offense and trial.

P is equal to one if the defendant has been previously convicted of one or more felonies and zero otherwise.

DRG is equal to one if the defendant is known to have illegally used controlled substances and zero otherwise.

S is equal to one if the defendant has been convicted of a sex offense and zero otherwise.

MEN is equal to one if the defendant has ever been retained in an institution for the mentally ill and zero otherwise.

AL is equal to one if the defendant is a known abuser of alcohol and zero otherwise.

Since the dependent variable is constrained to lie between zero and one, the logit function is used to estimate the parameters in equation 55.[3]

$$\ln\left\{\frac{P_B}{1 - P_B}\right\} = b_0 + b_1c_i + b_2c_iG + b_3G + b_4R + b_5A + b_6E + b_7L + b_8M + b_9D + b_{10}ED + b_{11}RE + b_{12}DEL + b_{13}P + b_{14}DRG + b_{15}S + b_{16}MEN + b_{17}AL + \epsilon + b_{18}c_i^2 + b_{19}c_i^2G \; . \tag{56}$$

Before presenting the results obtained from estimation of equation 56, results of related work by other authors will be briefly outlined.

## 6.2 Existing Empirical Work on the Probability of Bargaining

The results of three authors, William Rhodes,[4] Landes,[5] and Forst,[6] are presented as an aid in the interpretation and evaluation of the results of this study.[7]

William Rhodes used data for all criminal cases involving burglary, narcotics, larceny, and forgery during 1970 in Ramsey and Hennepin Counties, Minnesota. There were approximately 125 observations per category per county. OLS,[8], Log-Reciprocal,[9] and Probit Analysis[10] were used to analyze the data. In general, OLS and Probit estimated coefficients agreed in sign and magnitude while often Log-Reciprocal estimates differed substantially in magnitude and often in sign to the estimates of OLS and Probit Analysis. This divergence of results possibly stems from the manner in which zero values were handled in the Log-Reciprocal model. In this model ln(zero) was set to $-9 \times 10^{-6}$ for estimation purposes.[11] In his study bail was found to reduce the probability of bargaining in a limited number of cases.[12] However, if the defendant was released on personal recognizance the probability of bargaining fell for all types of crime. In all cases the probability of bargaining increased with age. Being a male or non-white also increased the probability of bargaining. Rhodes found no connection between the probability of bargaining and the judge who presided at trial.

In another study Landes used the data from an American Bar Study of over 11,000 defendants to estimate the probability of conviction. In this study the OLS estimation technique was used. Landes found the following variables to increase the probability of bargaining:

1. not receiving bail,
2. living in a metropolitan area,
3. being white,
4. low income,
5. expensive bail.

The only non-intuitive result is that being a white defendant increases the probability of bargaining. Landes dismisses this result since the coefficient with this variable is statistically insignificant.

Using 6,000 observations from the Washington, D.C., based computer information system PROMIS (Prosecutor's Management Information System) and OLS as the estimation technique, Forst and Brosi concluded that the following factors increase the probability of bargaining:

1. non-police witnesses,
2. being young,
3. existence of physical evidence,
4. committing a serious crime,
5. prior arrests,
6. robbing a business rather than a private residence,
7. unknown victim,
8. short time between offense date and trial,
9. few co-defendants.

## 6.3 Test for Direction of Shift in the Bargaining Space

If $b_2 = b_3 = b_{19} = 0$ or $\bar{c}b_2 + b_3 + \bar{c}^2 b_{19} = 0$ in equation 56, i.e., the gun law has no effect upon the probability of bargaining, then the second test described below is the last test that must be conducted in order to determine which case holds. If $b_2 = b_3 = b_{19} = 0$ or $\bar{c}b_2 + b_3 + \bar{c}^2 b_{19} = 0$ then only cases 1 or 4 can hold since in these two cases the probability of bargaining is not changed by imposition of the gun law. If one of the constraints on the parameters is satisfied and the bargained c falls as a result of the passage of the gun law, then case 1 must hold. This is the case generally described in Chapters III and IV. If the bargained c rises as a result of the passage of the gun law, then case 4 holds. If neither constraint on the parameters is satisfied, then a third test must be carried out. This test is outlined in Section 6.4 and in Chart 1 of that section.

The test (the second test referred to above) to determine the influence of the passage of the gun law on the bargained charge is carried out by estimation of a second logit function,

$$\ln\left\{\frac{c_i}{1 - c_i}\right\} = a_0 + a_1 G + a_2 B + a_3 S + a_4 A + a_5 E + a_6 R + a_7 L + a_8 M + a_9 c_f + a_{10} MEN + a_{11} DEL + a_{12} P + a_{13} ED + a_{14} DRG + a_{15} AL + a_{16} D + a_{17} RE \ . \quad (57)$$

$c_f$ is the Selling-Wolfgang measure of crime severity corresponding to the crime to which the defendant is ultimately convicted.[13]

$c_i$ is the Selling-Wolfgang measure of crime severity corresponding to the crime to which the defendant is initially charged.

B is equal to 1 if the defendant bargains and 0 otherwise.

All other variables are as previously defined.

If $a_1 < 0$ then passage of the gun law has reduced $c_f$, <u>ceteris paribus</u>, and case 1 holds; otherwise case 4 holds (given $b_2 = b_3 = 0$).

## 6.4 $b_2=b_3=b_{19}$ and $cb_2+b_3+c^2b_{19}$ Do Not Hold

If the probability of bargaining changes as a result of the passage of the gun law, it is necessary to determine if the new bargaining space is contained in, to the left or to the right of the pre-gun law bargaining space. If the restrictions on the b's described in Section 6.3 do not hold, then three rather than two tests must be performed to determine which case holds. This is so since when the probability of bargaining does not change as a result of passage of the gun law two possible cases are excluded.[14] However, when the probability of bargaining changes only one case is excluded by the first test. Case 3 is excluded if the probability of bargaining increases, while case 2 is excluded if the probability of bargaining falls as a result of the passage of the gun law.

The signing of coefficients in equation 2 is sufficient to exclude one more case. If c falls as a result of the passage of the gun law, then case 4 is excluded; however, if c rises case 1 is excluded.

Chart 1 is a schematic of the tests and their sequence of performance in order to determine which case holds. Also in the following chart $c_{GL}$ is the level of c after passage of the gun law and $c_{NGL}$ is the level of c, arrived at by bargain or trial, prior to

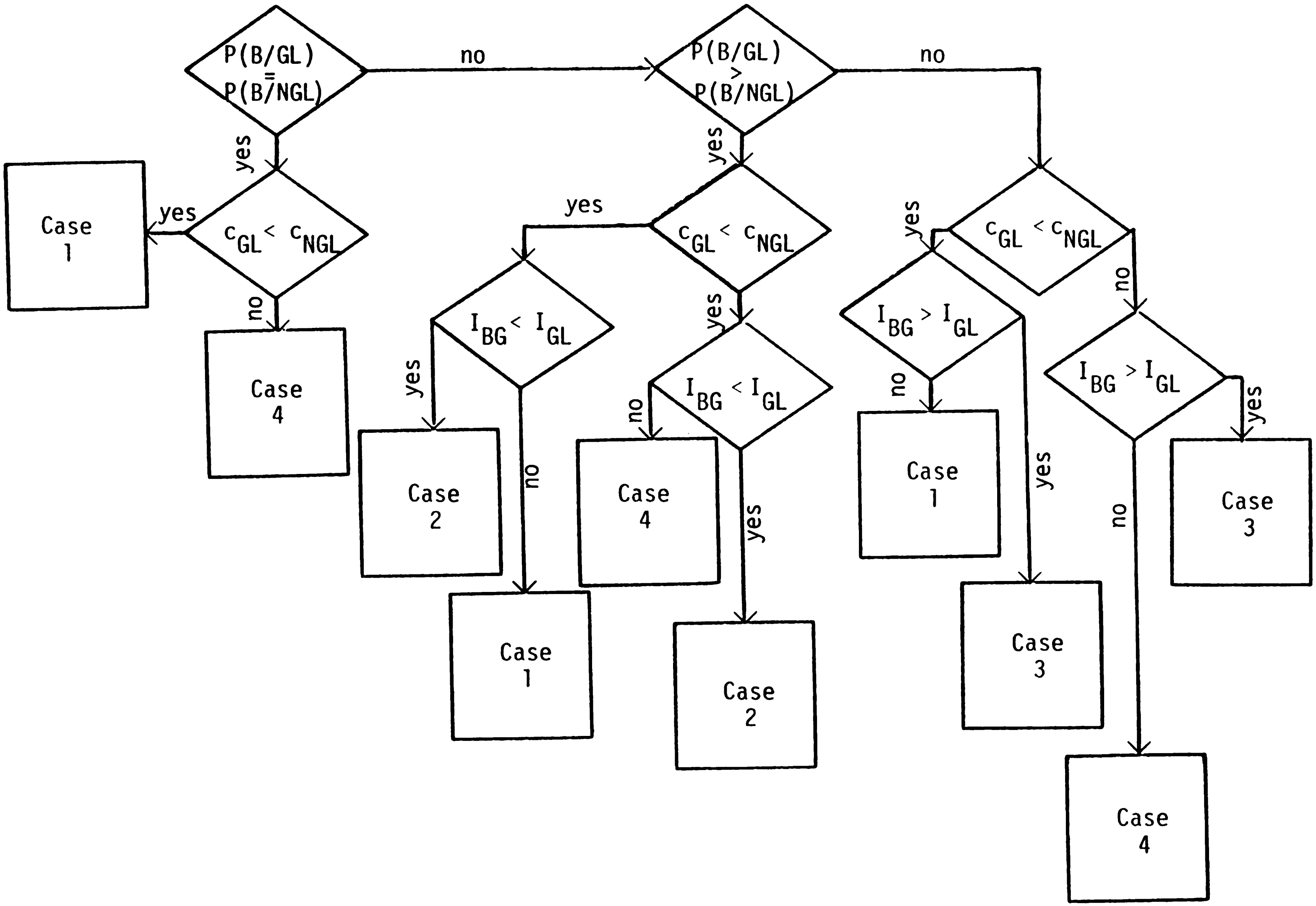



Chart 1.--Case selection.

passage of the gun law. P(B/GL) is the probability of bargaining given the gun law and P(B/NGL) is the probability of bargaining prior to passage of the gun law. $I_{BG}$ is the prediction interval of c before passage of the gun law and $I_{GL}$ is the prediction interval of c after the passage of the gun law.[15,16]

## 6.5 Results

Equation 56, the probability of bargaining, was estimated initially including all explanatory variables as outlined in Section 6.1.[17] As documented in Table 9, most explanatory variables were insignificant.

Table 9.--Results of initial estimation.

| Variable | Value | t |
|---|---|---|
| Gun | 1.158 | .464 |
| Age | .015 | .417 |
| Race | - .575 | -2.788 |
| Location | .466 | .628 |
| Lawyer | - .708 | - .937 |
| Education | - .084 | - .374 |
| Married | -1.890 | -2.349 |
| Employed | .422 | .546 |
| Related | .142 | .167 |
| Delay | - .002 | - .889 |
| Previous record | .874 | .133 |
| Drug | -1.020 | -1.490 |
| Sex crime | - .565 | - .421 |
| Mental | - .683 | - .706 |
| Alcohol | .525 | .459 |
| Charge*change | -14.130 | - .900 |
| Charge*gun | 26.380 | .967 |
| Charge*charge*gun | 6.460 | .289 |
| Charge | -25.248 | - .700 |
| Constant | 3.960 | 1.261 |

The two explanatory variables which were statistically significant (t's greater than 1.96) tended, given their presence, to reduce the probability of bargaining. Being non-Caucasian or married thus reduces the probability of bargaining. Table 10 compares the results presented in Table 9 with the work of the authors described in Section 6.2.

Table 10.--Comparison of results--full specification.

| Variable | Rhodes | Landes | Forst |
|---|---|---|---|
| Age | + | na | na |
| Race | na | + | na |
| Location | na | + | na |
| Related | na | na | - |
| Delay | na | na | + |
| Previous record | na | na | + |
| Charge | na | na | + |

same sign = +    opposite sign = 1    na = not applicable

A second logit estimation was performed which excluded all statistically non-significant variables except the measure for severity of charge. A Maximum Likelihood Ratio Test was carried out on the hypothesis that all statistically insignificant variables in the first regression were simultaneously equal to zero.[18] The $\chi^2_{15}$ of this test was 34.678. A value of this magnitude does not allow acceptance of the null hypothesis even at the 99.7 significance level. A third regression was performed which included the independent variables in equation 2 and two gun variables: the simple gun variable and the interaction term of charge and gun. In Table 11 the results of this regression are detailed.

Table 11.--Results of final estimation.

| Variable | Value | t |
|---|---|---|
| Gun | 2.468 | 2.47 |
| Race | - .731 | -1.32 |
| Married | -1.523 | -2.50 |
| Drug | -1.370 | -2.18 |
| Charge | 1.328 | .56 |
| Charge*gun | -7.566 | -2.20 |
| Constant | 1.458 | 2.04 |

Now both variables which include a gun term have statistically significant parameters. The total effect of the gun law on the probability of bargaining is 2.478*gun - 7.566*gun*charge. The appropriate statistical test was conducted with charge evaluated at its mean on the null hypothesis that 2.4668*gun - 7.566*gun*charge = 0.[19] At the 95 percent confidence interval it is impossible to reject the null hypothesis.[20] It appears then that the gun law has not significantly affected the probability of bargaining. The results of this test, however, do not strongly support the null hypothesis.

The second equation estimated was to measure the impact of the gun law on the level of concessions obtained by the defendant.[21] Two regressions were run. The first specification was that of the complete model as outlined in Section 6.3. The results of this formulation are presented in Table 12. The model then was re-estimated

Table 12.--Results of initial specification.

| Variable | B | Std. Error B | F / Significance | Beta / Elasticity |
|---|---|---|---|---|
| Gun | -.25517623 | .14308372 | 3.1805336<br>.079 | -.1246923<br>.05359 |
| Mental | .43955144E-01 | .20159256 | .47541229E-01<br>.828 | .0145578<br>-.00267 |
| Delay | .48741060E-03 | .49009619E-03 | .98907059<br>.324 | .0648421<br>-.03647 |
| Lawyer | .10066768 | .14945143 | .45371163<br>.503 | .0451871<br>-.03283 |
| Employed | -.77316334E-01 | .15445920 | .25056184<br>.618 | -.0328685<br>.00897 |
| Charge | 4.9935846 | 1.4445087 | 11.950457<br>.001 | .7858568<br>-.48298 |
| Married | .14503748E-01 | .17124567 | .71733380E-02<br>.933 | .0060648<br>-.00160 |
| Previous record | .10285894 | .14316584 | .51618493<br>.475 | .0493188<br>-.02899 |
| Education | .28746598E-01 | .41731599E-01 | .47450725<br>.493 | .0452104<br>-.13743 |
| Sex | .46548486E-01 | .27982163 | .27672510E-01<br>.868 | .0108127<br>-.00129 |
| Race | -.13523965 | .13792330 | .96146349<br>.330 | -.0662173<br>.03364 |

Table 12.--Continued.

| Variable | B | Std. Error B | F / Significance | Beta / Elasticity |
|---|---|---|---|---|
| Bargain | -.31158060 | .14256281 | 4.7767041 | -.1428626 |
| | | | .032 | .09816 |
| Drugs | -.72628010E-01 | .14577553 | .21461087 | -.0322117 |
| | | | .645 | .00963 |
| Alcohol | -.57388874E-01 | .22275918 | .66371890E-01 | -.0182462 |
| | | | .798 | .00317 |
| Location | .10889027 | .16140790 | .45512281 | .0488780 |
| | | | .502 | -.01505 |
| Related | -.32079937 | .16296750 | 3.8749405 | -.1470895 |
| | | | .053 | .04787 |
| Age | .18557680E-02 | .71161457E-02 | .68007638E-01 | .0209990 |
| | | | .795 | -.02389 |
| Offense | .64502852 | 1.3540904 | .22691465 | .1054302 |
| | | | .635 | -.06724 |
| Constant | -3.3933521 | .51684701 | 43.105615 | |

| | |
|---|---|
| Multiple R | .88189 |
| R Square | .77773 |
| Adjusted R Square | .71618 |
| Std. Deviation | .54590 |

including only those variables which were statistically significant in the first model specification. The results of this regression are presented in Table 13. Of special interest is that the gun coefficient is negative in both regressions. The null hypothesis that the gun law reduced the level of charge to which the defendant pleads guilty, ceteris paribus, cannot be rejected.

The data therefore do not reject the contention that Case 1 is the appropriate plea bargaining model.

Table 13.--Results of final estimation.

| Variable | B | Std. Error B | F / Significance | Beta / Elasticity |
|---|---|---|---|---|
| Charge | 5.6203403 | .38088327 | 217.74161<br>0 | .8844914<br>-.54361 |
| Gun | -.27918697 | .11478978 | 5.9153964<br>.017 | -.1364252<br>.05864 |
| Bargain | -.33657771 | .12061395 | 7.7870976<br>.007 | -.1543240<br>.10603 |
| Related | -.29844172 | .13195535 | 5.1152288<br>.026 | -.1368383<br>.04454 |
| Constant | -2.8742458 | .13315780 | 465.92347<br>0 | |

| | |
|---|---|
| Multiple R | .87296 |
| R Square | .76206 |
| Adjusted R Square | .75001 |
| Std. Deviation | .51233 |

## 6.6 Linear Probability Models

The simplest form of such a model[22] is:

1) $Y_i = b_0 + b_1 X_i + \epsilon_i$

where $Y_i = 1$ if the event of interest occurs and zero otherwise.

$X_i$ is a non-stochastic explanatory variable.

$\epsilon_i$ is an independently distributed random variable with zero mean.

If the expected value of equation 1 is taken then

$EY_i = b_0 + b_1X_i = P_1 x1 + (1-P_i)x0 = P_i$ where $P_i = P(Y_i = 1)$. Thus the expression in equation 1 may be interpreted as the probability that $Y_i$ occurs. It is easy to show that the error terms in equation 1 are heteroskedastic.[23] The estimated parameters in equation 1 are neither biased nor inconsistent because of the presence of heteroskedasticity.[24] The estimates, however, are inefficient in that they have larger sampling variances than those obtainable if the heteroskedasticity were removed. What is not normally mentioned is that the estimated covariance matrix is usually biased so that the standard tests of statistical significance are inappropriate. The bias in the covariance matrix is[25]

$$\sigma^2(X'X)^{-1}X'(I - V)X(X'X)^{-1} + \sigma^2 tr\frac{((X'X)^{-1}X'(I - V)X)(X'X)^{-1}}{(n - k)}$$

where $\sigma^2 V = E(\epsilon\epsilon')$. Theil[26] shows that heteroskedasticity need not lead to biasing of the variances. Theil shows this by using the simple model

$Y_i = b_0 + b_1X_i + {}_i$ where $\sigma_i^2 = E(\epsilon_i^2) = \gamma(E(Y_i))^2$ and $\sigma_b^2$ is the true variance of the slope coefficient and $s_b^2$ is the estimated variance of the slope coefficient. Then

$$2)\quad \sigma_b^2 = s_b^2\left\{\frac{1 + \dfrac{2rcu_3}{(u_2)^{3/2}} + r^2c^2\left\{\dfrac{u_4}{u_2} - 1\right\}}{1 + r^2c^2}\right\}$$

where r is the correlation coefficient, c is the coefficient of variation, and $u_i$ is the $i^{th}$ moment of X around its mean. The bias then depends upon the skewness and kurtosis of X, the correlation coefficient, and the coefficient of variation of Y.

Another point not stressed is that a linear specification of the model is a form of functional misspecification.[27] When the function to be fitted is misspecified, e.g., linear when in fact it is nonlinear, the parameter estimators are biased and inconsistent. Under certain conditions it can be shown that the slope parameters are not biased or inconsistent. However, the variances are still biased. It should be noted that use of linear restrictions does not remove the problem of misspecification[28] (unless the dependent variable is uniformly distributed).

The model can be viewed as being misspecified from another point. It was pointed out above that the error terms of equation 1 were heteroskedastic and thus the least squares estimators were inefficient. It is well known that the error terms in equation 1 have a discrete distribution;[29] in particular this implies that the error terms are not distributed normally. It is also well known that OLS is equivalent to MLE under the assumption that the error terms are normally distributed. Thus use of the specification in equation 1 as a probability model results in a form of model misspecification. This, as before, results in the estimators being biased and inconsistent.

Finally, it should be pointed out that $R^2$ is difficult to interpret in models that have binary dependent values.[30] It is possible for $R^2$ to be low yet for the model to be a perfect predictor. For example, consider the following problem:

| independent variable X | dependent variable Y | probability |
|---|---|---|
| 0 | 0 | .1 |
| 1 | 0 | .4 |
| 2 | 0 | .4 |
| 3 | 0 | .1 |

In this example Y can be predicted perfectly using the OLS estimate $Y = -.1 + 4X$ and the criteria $Y = 1$ when $Y = \frac{1}{2}$; otherwise $Y = 0$. Yet in this example $R^2 = .75$. Alternate measures of goodness of fit suggested are the correlation ratio, mean error probability, and average conditional entropy. None of these measures is without drawbacks. Some work has been done on estimating the upper bound to $R^2$ in problems with binary dependent variables. It has been shown that if Y is distributed, Beta (a,b) where the mean of the distribution is $\frac{a}{a + b}$ and the variance is $\frac{a \cdot b}{(a + b + 1)(a + b)^2}$ then the upper bound on $R^2$ is $\frac{1}{a + b + 1}$.[31] When a = b = 1 the Beta takes on the Uniform distribution. In this case the maximum $R^2$ is .33. The Beta distribution takes on the familiar Bell shape of the Normal distribution for a = b 2. This would imply a maximum $R^2$ of approximately .2.

## 6.7 The Logit Model

The logit model[32] is based upon the cumulative logistic probability function,

$$P_i = \frac{1}{1 + \exp(a + bX_i)} = \frac{1}{1 + Z} \tag{3}$$

where $P_i$ is the probability that the event of interest will occur and $X_i$ is a non-stochastic[33] variable which explains the occurrence (non-occurrence) of the event. An immediate problem is specification of the error term in equation 3. $P_i = \frac{1}{1+Z} + \epsilon_i$ is inappropriate if $\epsilon_i \sim N(0,\sigma^2)$ since this allows $P_i$ to fall outside the range of (0,1). $P_i = \frac{1}{1+Z+{}_i}$ keeps $p_i \in (0,1)$; however, $P_i$ is now not distributed $N(0,\sigma^2)$. Also there is no simple transformation which makes the model linear. However,

$$4) \qquad P_i = \frac{1}{1 + Z\exp(\epsilon_i)}$$

allows the model to be transformed to $\ln\left\{\frac{1-P_i}{P_i}\right\} = Z + \epsilon_i$. In this model $\ln\left\{\frac{1-P_i}{P_i}\right\} \sim N(0,\sigma^2)$. Note, however, that $P_i$ is not distributed normally.

## 6.8 Derivations of the Logistic Distribution

Several authors[34] have given theoretical justification for use of the logistic function. Let Y = 1 if E occurs and 0 otherwise. Also let X be a p vector with continuous density $h(X/E) \sim N(\theta_1,\Sigma_1)$, $h(X/\bar{E}) \sim N(\theta_2,\Sigma_2)$. Then by Bayes theorem

$$P(E/X) = \frac{P(E)h(X/E,\theta_1)}{P(E)h(X/E,\theta_1) + P(\bar{E})h(X/\bar{E},\theta_2)} = \frac{1}{1 + \frac{h(X,\bar{E})}{h(X,E)}} \qquad \text{since}$$

$$h(X/E,\theta_1) = \frac{1}{(2\pi)^{K/2}|\Sigma_1|^{\frac{1}{2}}} \exp(-\tfrac{1}{2}(X-\theta_1)'\Sigma_1^{-1}(X-\theta_1))$$

$$P(E/X) = \frac{1}{1 + \exp(-\frac{1}{2}(X-\theta_2)'\Sigma_2^{-1}(X-\theta_2) + \frac{1}{2}(X-\theta_1)'\Sigma_1^{-1}(X-\theta_1))} \quad .$$

Thus at worst P(E/X) is of the form $K\cdot\exp(X'AX + B'X + C)$ where A is a matrix, B is a row vector, and C is a constant. If $\Sigma_1^{-1} = \Sigma_2^{-1}$

then $P(E/X) = 1/\exp(-\frac{1}{2}(2(\theta_2'-\theta_1')\Sigma^{-1}X + (\theta_2-\theta_1)'\Sigma^{-1}(\theta_2-\theta_1)$ or $P(E/X)$ is of the form $1/\exp(X'b)$ as stated above.

In a purely statistical manner the logistic distribution is the limiting distribution of the mid-range (average of the largest and smallest values of a random sample of size n).[35] Let ${}_mY, Y_m$ be the $m^{th}$ observation in increasing and decreasing value. The mid-range is then $\frac{{}_1Y + Y_1}{2}$. It has been shown that the distribution of the $m^{th}$ mid-range tends toward normality for increasing m. The logistic distribution is also the limiting distribution of the extremal quotient $(Y_1/{}_1Y)$.[36] Finally, it has also been shown that if Y is distributed logistically then the distribution of $O_n$, the $n^{th}$ order statistic, is Weibull.[37] This result has been used by McFadden[38] to develop the logistic function from a set of axioms on consumer behavior.

## 6.9 The Relationship of the Logistic Distribution to the Normal Distribution

Because the logistic function takes on the general shape of the normal, the logistic has often been used in place of the normal because of its relative simplicity. Let the standard normal cumulative be denoted by $G_1$ and the standard logistic function be represented by $G_2$; then $G_2$-$G_1$ varies as shown in Figure 52 on the following page.[39] The maximum value of $G_2$-$G_1$ is .0228 and is attained at X = .7. The maximum .0228 may be reduced to less than .01 by transforming X in $G_1$ to (16/15)X. The new difference $G_2$-$G_1((16/15)X)$ is represented by the dashed curve in Figure 52.

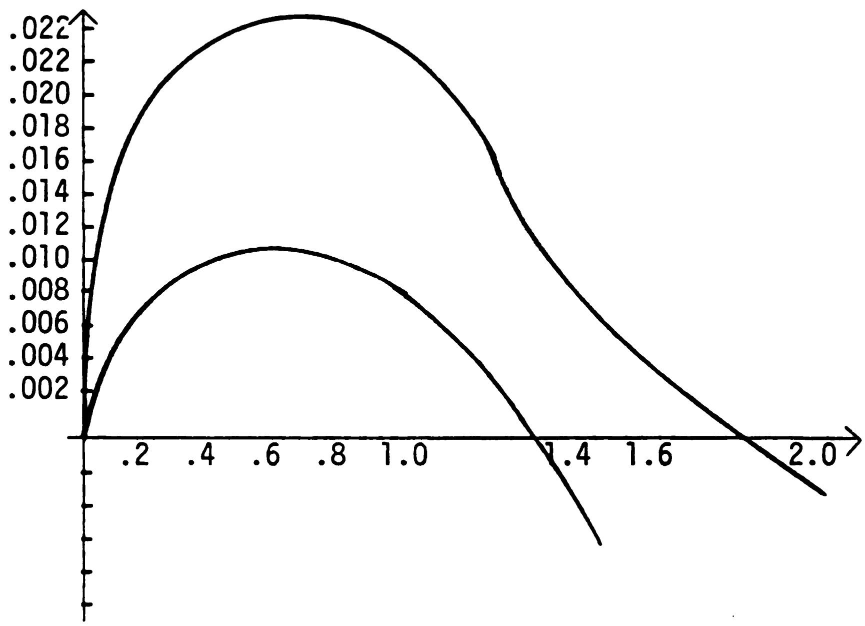



Figure 52.--Difference of logistic and normal distributions.

The kurtosis of the logistic is 4.2 and that of the normal is 3. This difference results from the relatively long tails of the logistic distribution. However, Plackett[40] using the maximum likelihood estimation of the parameters of the logistic distribution obtained results quite similar to those of the BLUE[41] of the normal distribution, even for samples of size no greater than 10.

## 6.10 Estimation of the Logit

Transforming equation 3 to

5) $$\ln\left\{\frac{1 - P_1}{P_i}\right\} = a + bX_i + \epsilon_i$$

cannot be directly estimated if either $P_i = 0$ or 1 since then the right hand size of 5 is undefined. If there are sufficient observations the data may be grouped where group i has $n_i$ observations with $r_i$ having the desired trait. Thus an estimate of $P_i = P_i = r_i/n_i$ is obtained. Instead of estimating equation 5 above, the following is estimated:

$$6) \quad \ln\left\{\frac{n_i - r_i}{n_i}\right\} = a + bX_i + \epsilon_i.$$

For small samples the rhs of equation 6 is not normally distributed.[42] Thus OLS will produce consistent estimators only when the sample size is large since the distribution of the rhs of 6 is normally distributed. Also since equation 6 involves grouped data, efficiency can be improved by correcting for heteroskedasticity.

Direct estimation of equation 3 using MLE results in an estimator which is consistent,[43] asymptotically efficient, and asymptptically normally distributed.

## 6.11 Data Summary

Summary statistics for all variables used in this study are presented in this section. These statistics include for each variable: mean, variance, standard deviation, skewness, kurtosis, minimum, maximum, sum, percentage coefficient of variation, and the 95 percent confidence interval.

Variable: Previous Convictions

| | | | | | |
|---|---|---|---|---|---|
| Mean | .607 | Std. err | .054 | Std. dev. | .491 |
| Variance | .241 | Kurtosis | -1.845 | Skewness | -.447 |
| Minimum | 0 | Maximum | 1.000 | Sum | 51.000 |

Variable: Drugs

| | | | | | |
|---|---|---|---|---|---|
| Mean | .286 | Std. err. | .050 | Std. dev. | .454 |
| Variance | .207 | Kurtosis | -1.093 | Skewness | .966 |
| Minimum | 0 | Maximum | 1.000 | Sum | 24.000 |

Variable: Sex Offender

| | | | | | |
|---|---|---|---|---|---|
| Mean | .060 | Std. err. | .026 | Std. dev. | .238 |
| Variance | .057 | Kurtosis | 12.676 | Skewness | 3.791 |
| Minimum | 0 | Maximum | 1.000 | Sum | 5.000 |

Variable: Mental Patient

| | | | | | |
|---|---|---|---|---|---|
| Mean | .131 | Std. err. | .037 | Std. dev. | .339 |
| Variance | .115 | Kurtosis | 3.035 | Skewness | 2.228 |
| Minimum | 0 | Maximum | 1.000 | Sum | 11.000 |

Variable: Alcohol Abuse

| | | | | | |
|---|---|---|---|---|---|
| Mean | .119 | Std. err. | .036 | Std. dev. | .326 |
| Variance | .106 | Kurtosis | 3.830 | Skewness | 2.396 |
| Minimum | 0 | Maximum | 1.000 | Sum | 10.000 |

Variable: Offense

| | | | | | |
|---|---|---|---|---|---|
| Mean | .225 | Std. err. | .018 | Std. dev. | .167 |
| Variance | .028 | Kurtosis | .542 | Skewness | 1.393 |
| Minimum | .040 | Maximum | .600 | Sum | 18.860 |

Variable: Lawyer

| | | | | | |
|---|---|---|---|---|---|
| Mean | .702 | Std. err. | .050 | Std. dev. | .460 |
| Variance | .212 | Kurtosis | -1.217 | Skewness | -.901 |
| Minimum | 0 | Maximum | 1.000 | Sum | 59.000 |

Variable: Education Level

| | | | | | |
|---|---|---|---|---|---|
| Mean | 10.298 | Std. err. | .176 | Std. dev. | 1.612 |
| Variance | 2.597 | Kurtosis | 4.080 | Skewness | -.925 |
| Minimum | 3.000 | Maximum | 14.000 | Sum | 865.000 |

Variable: Married

| | | | | | |
|---|---|---|---|---|---|
| Mean | .238 | Std. err. | .047 | Std. dev. | .428 |
| Variance | .184 | Kurtosis | -.443 | Skewness | 1.252 |
| Minimum | 0 | Maximum | 1.000 | Sum | 20.000 |

Variable: Employed

| | | | | | |
|---|---|---|---|---|---|
| Mean | .250 | Std. err. | .048 | Std. dev. | .436 |
| Variance | .190 | Kurtosis | -.633 | Skewness | 1.176 |
| Minimum | 0 | Maximum | 1.000 | Sum | 21.000 |

Variable: Related

| | | | | | |
|---|---|---|---|---|---|
| Mean | .321 | Std. err. | .051 | Std. dev. | .470 |
| Variance | .221 | Kurtosis | -1.428 | Skewness | .779 |
| Minimum | 0 | Maximum | 1.000 | Sum | 27.000 |

Variable: Delay

| | | | | | |
|---|---|---|---|---|---|
| Mean | 161.179 | Std. err. | 14.874 | Std. dev. | 136.318 |
| Variance | 18582.703 | Kurtosis | 9.765 | Skewness | 2.443 |
| Minimum | 20.000 | Maximum | 897.000 | Sum | 13539.000 |

Variable: Gun

| | | | | | |
|---|---|---|---|---|---|
| Mean | .452 | Std. err. | .055 | Std. dev. | .501 |
| Variance | .251 | Kurtosis | -2.010 | Skewness | .195 |
| Minimum | 0 | Maximum | 1.000 | Sum | 38.000 |

Variable: Bargain

| | | | | | |
|---|---|---|---|---|---|
| Mean | .679 | Std. err. | .051 | Std. dev. | .470 |
| Variance | .221 | Kurtosis | -1.428 | Skewness | -.779 |
| Minimum | 0 | Maximum | 1.000 | Sum | 57.000 |

Variable: Age

| | | | | | |
|---|---|---|---|---|---|
| Mean | 27.726 | Std. err. | 1.265 | Std. dev. | 11.595 |
| Variance | 134.442 | Kurtosis | 4.427 | Skewness | 2.110 |
| Minimum | 17.000 | Maximum | 74.000 | Sum | 2329.000 |

Variable: Race

| | | | | | |
|---|---|---|---|---|---|
| Mean | .536 | Std. err. | .055 | Std. dev. | 5.02 |
| Variance | .252 | Kurtosis | -2.028 | Skewness | -.146 |
| Minimum | 0 | Maximum | 1.000 | Sum | 45.000 |

Variable: Gender

| | | | | | |
|---|---|---|---|---|---|
| Mean | 1.000 | Std. err. | 0 | Std. dev. | 0 |
| Variance | 0 | Kurtosis | 0 | Skewness | 0 |
| Minimum | 1.000 | Maximum | 1.000 | Sum | 84.000 |

Variable: Location

| | | | | | |
|---|---|---|---|---|---|
| Mean | .298 | Std. err. | .050 | Std. dev. | .460 |
| Variance | .212 | Kurtosis | -1.217 | Skewness | .901 |
| Minimum | 0 | Maximum | 1.000 | Sum | 25.000 |

Variable: Charge

| | | | | | |
|---|---|---|---|---|---|
| Mean | .208 | Std. err. | .018 | Std. dev. | .161 |
| Variance | .026 | Kurtosis | 1.013 | Skewness | 1.533 |
| Minimum | .040 | Maximum | .600 | Sum | 17.500 |

Variable: Conviction

| | | | | | |
|---|---|---|---|---|---|
| Mean | .142 | Std. err. | .016 | Std. dev. | .144 |
| Variance | .021 | Kurtosis | 3.383 | Skewness | 2.001 |
| Minimum | .020 | Maximum | .600 | Sum | 11.940 |

Footnotes--Chapter VI

[1]See Chapter IV, Sections 11.4 and 11.5.

[2]The measure of crime severity was constructed using the procedure detailed in Sellin Thorsten and Marvin E. Wolfgang, The Measurement of Delinquency (New York: John Wiley and Sons, 1964). In particular, see Chapter 17, "Deriving Scores," pp. 274-91.

[3]See Section 6.6 for details of logit specification.

[4]William Rhodes, "A Study of Sentencing in the Hennepin County and Ramsey County District Courts," The Journal of Legal Studies 5 (June 1976):333-453.

[5]William Landes, "An Economic Analysis of the Courts," Journal of Law and Economics 14 (April 1971): 61-108.

[6]Brian Forst and Kathleen B. Brosi, "A Theoretical and Empirical Analysis of the Prosecutor," The Journal of Legal Studies 6 (January 1977): 177-91.

[7]A measure of relative strength of defendant to prosecutor has been developed from information obtained from the Detroit Recorders Court by J. A. Lachman, "An Economic Model of Plea Bargaining in the Criminal Court System" (Ph.D. dissertation, Michigan State University, 1975). It is based upon the prosecutor's estimate of the probability of getting defendant X convicted. Its exact form is:

$$\ln\left\{\frac{1-r}{r}\right\} = 1.13 - \underset{(3.55)}{1.2\ \text{Jail}} - \underset{(-1.98)}{.55\ \text{Race}} + \underset{(2.63)}{1.05\ \text{Race}\cdot\text{Jail}}$$
$$+ \underset{(1.07)}{.002\ \text{Total Time}} - \underset{(-1.94)}{.35\ \text{Pend Chg}} + \underset{(1.16)}{.3\ \text{Know Wit}}$$

Where Jail equals 1 if the defendant is in custody (no bail) and 0 otherwise,
Race equals 1 if the defendant is non-Caucasian and 0 otherwise,
Race·Jail equals 1 if the defendant is non-Caucasian and in jail and 0 otherwise,
Total Time is the number of days between the offense and trial,
Pend Chg is 1 if the defendant has other cases pending and 0 otherwise,
Know Wit is 1 if the defendant and the complainant are friends or relatives and 0 otherwise.

Lachman was not the first to construct such a measure, but her use of the measure was original. For an example of an earlier construction of such a measure see William M. Landes, "An Economic Analysis of the Courts," Journal of Law and Economics 14 (April 1971): 61-108. More recently, Brian Forst and Kathleen Brosi, "A Theoretical and Empirical

Analysis of the Prosecutor," The Journal of Legal Studies 6 (January 1977): 177-92 have constructed an index similar to that of Lachman using OLS. The magnitude of the coefficients differs from those in Lachman's study but the signs are the same in general. This work of Lachman is not presented in the main body of this dissertation since it is not clear how to interpret the signs of the coefficients presented above. Lachman interprets the coefficients as if the dependent variable were written ln(r/(1-r)) but she writes the equation as ln((1-r)/r). If the latter form was used in estimation, then the signs on the coefficients should be changed before interpretation. Since not all signs were as expected, it is impossible to tell which form of the logit specification was actually estimated.

[8]For a discussion of Ordinary Least Squares (OLS) see Jan Kmenta, Elements of Econometrics (New York: The Macmillan Co., 1971), pp. 197-306.

[9]J. Johnston, Econometric Methods, 2nd ed. (New York: McGraw-Hill Book Co., 1972), pp. 52-53.

[10]D. J. Finney, Probit Analysis, erd ed. (Cambridge: Cambridge University Press, 1971).

[11]Rhodes, p. 339.

[12]In this as the other cases cited, the probability of bargaining was not directly analyzed. Rhodes and others examined the probability of receiving a prison sentence. From these results inferences are drawn based upon the theory presented in 2.1.3.

[13]This second logit function need not be estimated by MLS since values of $C_f$ exist. Thus equation 2 can be estimated via OLS since $\ln\left\{\frac{1-C_f}{C_f}\right\}$ is always well defined.

[14]Cases 2 and 3 are excluded.

[15]Let A be a 9x1 column vector of the estimated parameters in equation 2. Let C* be a value of c not necessarily used in estimation of A. Let X*, a row vector, be those values of the explanatory values which product C*. Then the

$$P[X^* A' - t\alpha/_2 \cdot S\sqrt{1+a'\ (x'x)^{-1}A} \leq C^*$$

$$\leq XA' + t\alpha/_2\ S\sqrt{1 + A'\ (x'x)^{-1}A}\ ] = 1 - \alpha\ .$$

This interval is different from a confidence interval in the following way. In a confidence interval bounds are found for a fixed parameter. In a prediction interval, not only are the bounds random but also the parameter (C*) for which bounds are sought. See Henri Theil, Principles of Economics (New York: John Wiley and Sons, 1971), 134-35.

[16]For a complete discussion of decision trees and their applications see Howard Raiffa, Decision Analysis (Reading, Pa.: Addison-Wesley, 1976).

[17]The probability of bargaining is equal to 1 if a bargain was struck and 0 otherwise.

[18]Alexander M. Mood, Franklin A. Graybill, and Duane C. Boes, Introduction to the Theory of Statistics, 3rd ed. (New York: McGraw-Hill, 1974), p. 440.

[19]Ibid., p. 435 (a composite t test).

[20]The null hypothesis is that 2.4668Gun - 7.566Gun*charge = 0.

[21]This time the logit estimation did not require the use of a Maximum Likelihood estimation technique since $c_i$ is not a binary variable. This fact allows the use of Ordinary Least Squares.

[22]Kmenta, pp. 425-28.

[23]Ibid., p. 426.

[24]Ibid., pp. 250-52.

[25]Stephen Goldfeld and Richard Quandt, Nonlinear Methods in Econometrics (New Holland: North Holland Publishing Co., 1972), p. 81.

[26]H. Theil, "Estimates and Their Sampling Variance of Parameters of Certain Heteroskedastic Distribution," Review de L'Institut International de Statistique 19 (1951): 141-47.

[27]Kmenta, Chapter 12, section 4, pp. 392-402.

[28]R. S. Pindyck and D. L. Rubinfeld, Econometric Models and Economic Forecasts (New York: McGraw-Hill, 1976), pp. 242-43.

[29]Ibid., p. 240.

[30]The following example is taken from John Neter and E. S. Maynes, "On the Appropriateness of the Correlation Coefficient with a 0,1 Dependent Variable," Journal of the American Statistical Association 65 (June 1970): 501-509.

[31]D. Morrison, "Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions," Journal of the American Statistical Association 67 (March 1972): 68-70.

[32]N. Johnson and Samuel Kotz, Distributions in Statistics, Continuous Univariate Distributions, Vol. 2 (New York: Houghton Mifflin Col, 1972), Chapter 22.

[33]It is not necessary for $X_i$ to be nonstochastic but then certain other conditions must be satisfied. See Kmenta, pp. 297-304.

[34]Stanley Warner, "Multivariate Regression of Dummy Variates Under Normality Assumptions," Journal of the American Statistical Association 58 (December 1973): 1054-63.

[35]E. J. Gumbel, "Ranges and MidRanges," Annals of Mathematical Statistics 15 (1944): 414-22.

[36]E. J. Gumbel and D. R. Keeney, "The External Quotient," Annals of Mathematical Statistics 21 (1950): 523-38.

[37]Mood, Graybill and Boes, pp. 259-60.

[38]D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers of Econometrics, ed. Paul Zarembka (New York: Academic Press, 1974), pp. 105-42.

[39]Johnson and Kotz, pp. 5-6.

[40]R. L. Plackett, "Linear Estimation from Censored Data," Annals of Mathematical Statistics 29 (1958): 131-42.

[41]Best Linear Unbiased Estimator.

[42]The $r_i$ are binomially distributed.

[43]Thiel, pp. 392-96.

## CHAPTER VII

## CONCLUSIONS

In this dissertation a stochastic model of the bargaining behavior of the prosecutor and defendant was developed. Analysis focused on the effect of the introduction of a mandatory minimum sentence upon the boundaries of the bargaining space. These bargaining boundaries were defined by the construction of switch functions.[1] A theory was developed which viewed the introduction of a minimum sentence as a change in the probability density function associated with the prison term to be received upon conviction for a given charge. The theory implied that the imposition of the minimum sentence would result in the switch functions of the prosecutor and defendant shifting and that this shifting of the switch functions would impact upon the plea negotiation process. Logically this impact could be broken down into four cases.[2] Tests were developed to discriminate among the four cases. It was found that the only case consistent with the data was Case 1.[3] In Case 1 the introduction of a mandatory sentence resulted in the prosecutor's willingness to reduce the level of charge stemming from bargaining, *ceteris paribus*, to increase. At the same time the introduction of the mandatory sentence law resulted in the defendant's increased efforts to reduce the level of the plea negotiated charge.[4] This would suggest that if the intent of the legislature was to increase the length of time spent in prison by those defendants convicted of crimes carrying the mandatory sentence that their goal has been in large unfulfilled.[5]

The prosecutor by reducing the normal plea bargain charge for a given offense, after passage of the gun law, has accomplished at least two things. First, it allowed him to maintain a high-level percentage of cases resolved via plea negotiation in the face of stiffer bargaining by the defendant. Second, it allows the judiciary more freedom in the actual sentencing process. These results would support the findings of Rhodes that bargaining and severity of charge that the defendant is convicted for in relation to the actual offense is closely related to the case load per prosecutor.[6] If the intent of the legislature is to increase the length of sentence it most likely cannot be accomplished simply by fiat. Their intent will be frustrated by the prosecutor, judiciary, and the parole department unless backed up by adequate financing.[7] Additional funding is essential since mandatory sentencing if faithfully carried out would increase the work load of all agencies involved. The prosecutor's case load would increase since a larger number of defendants would refuse to bargain, resulting in increased trial preparation by the prosecutor's department. The judiciary in addition to losing sentencing flexibility would experience an increase in work load since more cases would be coming to trial. The parole and penology system would have to determine what to do with the extra inmates since the length of average stay would go up. Thus there are incentives at every level encouraging the failure of mandatory sentencing laws.

The conclusions of this dissertation lead to the need for additional research to verify its findings, since it implies a shift in legislative policy. First the signs of dc/dr and dV/dr should be directly estimated. In the chapters on the prosecutor and defendant, the Trans-Log Utility Functions for the prosecutor and defendant were developed and the signing of parameters necessary for determining dc/dr and dV/dr identified. Direct estimation of the Trans-Log Utility Function would require the extensive cooperation of a prosecutor's office, since the cost and defendant data required are not normally generated. Second, an attempt could be made to construct the prosecutor's and defendant's utility functions through interview techniques. These functions could then be examined for consistency with the results of the empirical work of this dissertation. This interview approach has been receiving increasing attention in recent years.[8]

## Footnotes--Chapter VII

[1]See Sections 3.10 and 3.11 for more on the switch function.

[2]See Section 4.16 for a description of the four cases.

[3]See Sections 4.17, 6.3, and 6.4 for a description of the tests.

[4]The selection of Case 1 depended in part upon the gun law having no effect upon the probability of bargaining. This was found to hold since $b_1G + b_3Gc + b_{19}Gc^2$ when evaluated at the mean of c (.208) was found to be zero. The following table gives the t values for the hypothesis that $b_1G + b_3Gc = 0$ for different values of c.

Table 14.--Ho: $b_1G + b_3Gc = 0$.

| c | t |
|---|---|
| 0.00 | 2.47 |
| 0.10 | 2.28 |
| 0.20 | 1.59 |
| 0.208 | 1.50 |
| 0.30 | 0.32 |
| 0.40 | -0.68 |
| 0.50 | -1.21 |
| 0.60 | -1.49 |
| 0.70 | -1.66 |
| 0.80 | -1.76 |
| 0.90 | -1.84 |
| 1.00 | -1.89 |

For values of c around the mean the introduction of the gun law has not affected the probability of bargaining. For low values of c, however, it appears that the introduction of the gun law has increased the probability of bargaining, while for high values of c the introduction of the gun law has reduced the probability of bargaining. Chart 1 of Section 6.4 shows that if the probability of bargaining increases as a result of the introduction of the gun law, *ceteris paribus*, then either Case 1 or Case 2 may hold. Case 2 implies the same behavior as Case 1 for the prosecutor. But the defendant now is willing to accept a higher charge than before. This was not looked at within this dissertation but possibly a preponderance of first-time offenders are associated with low c's and they respond differently than the other defendants. For defendants with high charges, the probability of bargaining falls because of the passage of the gun law. Referring again to Chart 1 of Section 6.4, it can be seen that Cases 1 and 3 are possible. This time it is not the defendant whose behavior switches but rather the prosecutor's. In this situation it might happen that the prosecutor will seek a higher charge than before passage of the gun law for a given charge. This could possibly be a reaction to exposure. High charges are associated with violent crimes -- the types of crimes that get public attention. This attention could possibly account for the shift in the prosecutor's behavior.

[5]Table 6 and supporting narrative in Chapter V reveal that little difference exists in the length of prison sentence received by defendants convicted under the gun law relative to defendants convicted of similar crimes prior to the mandatory sentencing law.

[6]William Rhodes, "The Economics of Criminal Courts: A Theoretical and Empirical Investigation," The Journal of Legal Studies 5 (June 1976): 333.

[7]The data used in this study were collected by the parole department of Michigan. The intent of this data-collection effort was to determine the impact of the mandatory sentencing law on the length of incarceration.

[8]David E. Bell, Ralph L. Keeney, and Howard Raiffa, eds., Conflicting Objectives in Decisions (New York: John Wiley and Sons, 1977). See also Ralph L. Keeney and Howard Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs (New York: John Wiley and Sons, 1976).

APPENDIX A

# APPENDIX A

The concepts of risk and changes in risk are used extensively in the development of the models for prosecutor and defendant. In this appendix these notions of risk will be presented, along with an outline of their historical development.

Risk, as used in this study, was initially described in terms of stochastic dominance[1] and efficiency criteria,[2] different names for the same idea.[3] Most work involving these concepts has centered on analysis of financial markets.[4] It is only recently that efforts to extend use of stochastic analysis beyond financial markets have become prevalent.[5]

The order of presentation in this appendix will be as follows:

1. concept of stochastic dominance as developed by Hadar, Hanoch, and Levy;
2. extensions made by Stiglitz and Rothschild,
3. discussion of how the Stiglitz and Rothschild results may be further extended; and
4. extensions to the theory by J. Meyer which incorporate the decision maker's level of risk aversion into the analysis of ranking of levels of risk.

## A.1 Approaches to Risk Analysis

The decision maker's utility function is an abstraction from reality. In general, because of difficulties of measurement[6] and a belief that a general specification will not be adequate to represent different individuals, economists in general seek to undertake analysis that does not depend upon knowledge of the specific form of the utility function. Thus, analysis that includes the utility function has sought to minimize the importance of its inclusion upon the results obtained. Of course, some assumptions about the form of the utility function must be made. These relate to the risk-averse nature (or lack of it) of the decision maker or class of decision makers under analysis. The general restrictions are that the utility function is concave or convex in all variables. If conclusions are to be drawn about movements in variables because of changes in risk, then the convexity (or concavity) of the marginal utilities, in the stochastic variable, must be specified. It should be noted that none of these restrictions requires that the utility function be given a specific functional form.

Because of the general nature of the assumptions, it is not always possible to assess the impact of changes in risk upon utility. Recent work by Meyer, discussed below, describes a technique for partially overcoming this problem of indeterminacy.

## A.2 The Initial Work of Hanoch, Levy, Russell, and Hadar[7]

The question Hanoch, Levy, Russell, and Hadar sought to answer is this: Assume there are two possible states to the world and that the random variable X has associated with it a different cumulative frequency distribution in each state of the world. For example, let Y represent the time defendant X will spend in jail if convicted. Y is a random variable with a cumulative distribution F, which portrays the beliefs of the prosecutor as to the likelihood of any particular sentence being served by defendant X. Let G represent the distribution of Y given that defendant X is convicted of a charge having a mandatory minimum sentence. The assumption is that there exist sentences feasible under F that are not feasible under G. That is, there are sentences allowable under F that are below the minimum that exists in the state of the world associated with G. Let U represent the utility function of the prosecutor and that he seeks to maximize his expected utility. The problem the above authors sought to solve is:

$$\text{When is } \int_a^b U(Y)dF(Y) \geq \int_a^b U(Y)dG(Y)?$$

That is, when does the prosecutor seek a conviction that does not carry a mandatory sentence? In the above, a is the shortest feasible sentence and b is the longest sentence that has a chance of occurring. Before continuing, it is necessary to develop some terminology.

## A.3 Stochastic Dominance

Let $U_1$ be the set of all utility functions that are continuous, bounded, and monotonically increasing on [a,b], where [a,b] is a portion of the real line. Let $U_2$ be the set of all utility functions that are continuous, bounded and $U'' < 0$ on [a,b]. Let $U_3$ be all those utility functions in $U_2$ where $U''' > 0$. This partial ordering of utility functions may be written more compactly in the following manner:

$U_1$ = {U: U, U' are continuous and bounded on [a,b] and $U' > 0$},

$U_2$ = {U: $U \in U_1$, U" is continuous and bounded on [a,b] and

$U'' < 0$ on [a,b] }, and

$U_3$ {U: $U \in U_2$, U" is continuous and bounded on [a,b] and

$U''' > 0$ on [a,b] }.

Given these definitions of $U_i$ , i=1, 2, or 3, it is possible to define stochastic dominance.

Definition 1: Stochastic Dominance

$F >_1 G$ if and only if $\int U(X)dF(X) > \int U(X)dG(X)$ for all $U \in U_i$

where $>_1$ ≡ first degree stochastic dominance,

$>_2$ ≡ second degree stochastic dominance, and

$>_3$ ≡ third degree stochastic dominance.

<u>Some Practical Properties</u>. Since $U_{J-1}$ is contained in $U_J$ J = 2,3 , it follows that dominance in $U_J$ implies dominance in $U_K$ for all K less than J. Thus, if $F >_1 G$ then $F >_2 G$ and $F >_3 G$.

Definition 2: Partial Ordering

Binary Relation. R is a binary relation defined on (the set) S if, for every pair of elements X, Y $\in$ S, where the ordering of X,Y is crucial, the statement XRY is either true or false.

A binary relation leads to a quasi ordering or preordering if the following two properties hold:

a. $XRS \; \forall X \in S$ (reflexivity) and

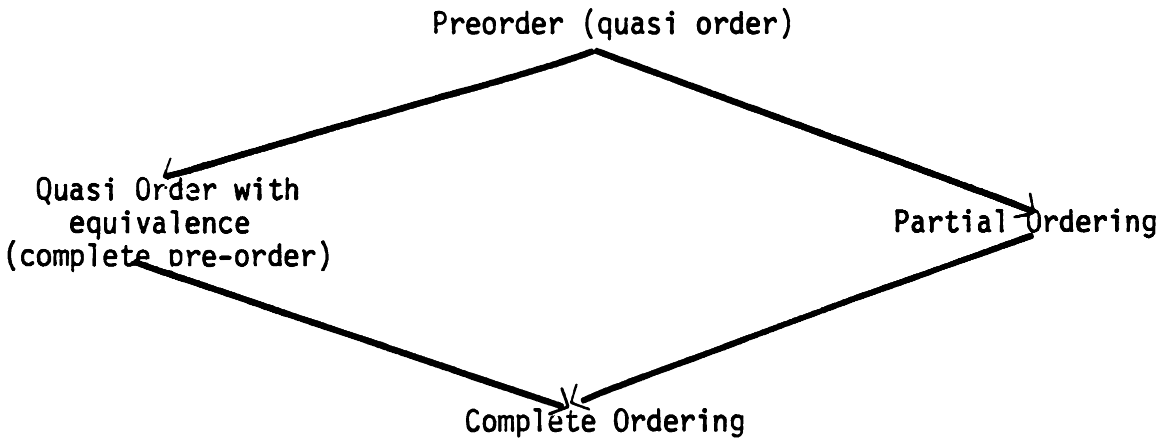
b. $XRY$ and $YRZ \Rightarrow XRZ$ (transitivity).

If the preordering (quasi ordering) defined above also has the following property:

c. $XRY$ and $YRX \Rightarrow X = Y$ anti-symmetric (asymmetric),

then the preordering defined by R is a partial ordering. If both XRY and YRX but $X \neq Y$, then X and Y are said to be equivalent and R is still a quasi ordering. The quasi ordering or preordering is complete if XRY and/or YRX for all $X, Y \in S$.

Thus the binary relation R can induce the following types of orders.[8]




This may be rewritten with the associated properties as:

Preorder
i. reflexivity
ii. transitivity

Quasi Order
(complete pre-order)
i. reflexivity
ii. transitivity
iii. equivalence

Partial Ordering
i. reflexivity
ii. transitivity
iii. anti-symmetry

Complete Ordering
i. reflexivity
ii. transitivity
iii. equivalence or antisymmetry
iv. XRY or YRX for all X, Y $\in$ S

where an arrow indicates decreasing generality.

Three properties which the three dominance orderings have in common are: (i) asymmetry, (ii) transitivity, and (iii) reflexivity. This implies that $>_i$ : , $i = 1, \ldots, 3$ are partial orderings over risky prospects.

## A.4 Determination of Stochastic Dominance

In development of necessary and sufficient conditions for stochastic dominance, it is useful to note that for all U(X) non-decreasing and bounded,

$$\int U(X)dF(X) - \int U(X)dG(X)$$
$$= \int U(X)\,[dF(X) - dG(X)]$$
$$= \int [G(X) - F(X)]\,(dU(X))$$
$$= \int S(X)dU(X)$$

where

$$S(X) = G(X) - F(X)^{9}$$

## A.5 First Order Stochastic Dominance

For all $U \in U_1$, a necessary and sufficient condition for $\int U(X)dF(X) \geq \int U(X)dG(X)$ is:

a. $F(X) \leqq G(X) \;\forall X \in [a,b]$ and

b. $F(Xo) < G(Xo)$ for at least one $Xo \in [a,b]$.

## A.6 Second Order Stochastic Dominance

A necessary and sufficient condition for $\int U(X)dF(X) \geq \int U(X)dG(X)$ for all $U \in U_2$ is:

$$\int_a^Y S(X)dX \geq 0; \;\forall Y \in [a,b],$$

where $S(X) = G(X) - F(X)$.

## A.7 Difference in First and Second Order Stochastic Dominance

In the case of first degree stochastic dominance, it is required that $F(X) \leq G(X)$ for all X. If we are concerned in ranking risky prospects (differentiated from one another by different cumulative distribution), then only a limited number of risky prospects can be ordered; that is, it is only possible to distinguish between those prospects for whom the cumulative distribution functions do not cross.

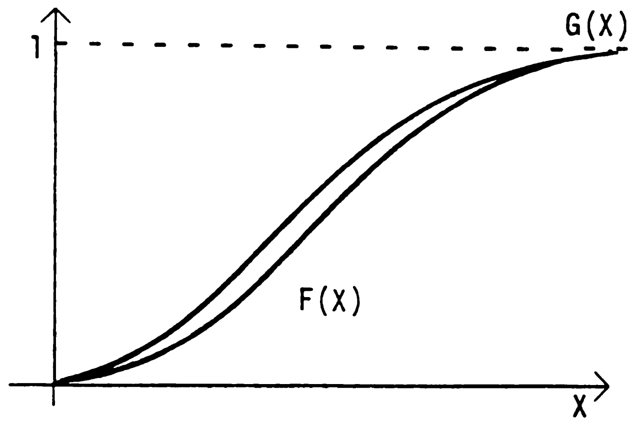



Figure 53.--Case where F dominates G under first order stochastic dominance for all U in $U_1$.

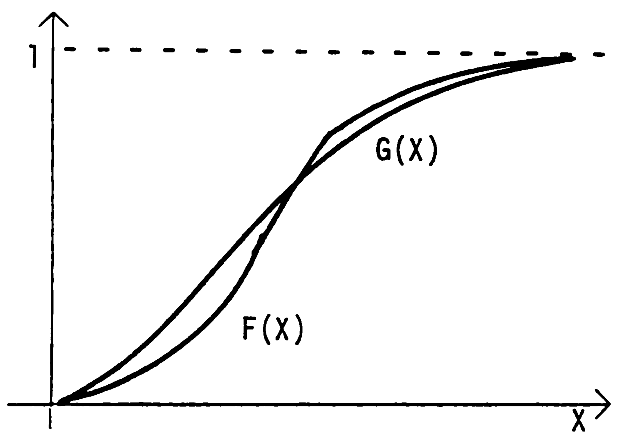



Figure 54.--Case where F does not dominate G under first order stochastic dominance.

If the additional constraint U" $\leq$ 0 is included in addition to the specification that the utility function be monotonically increasing, then it is possible to order some risky prospects whose cumulative distributions do cross.

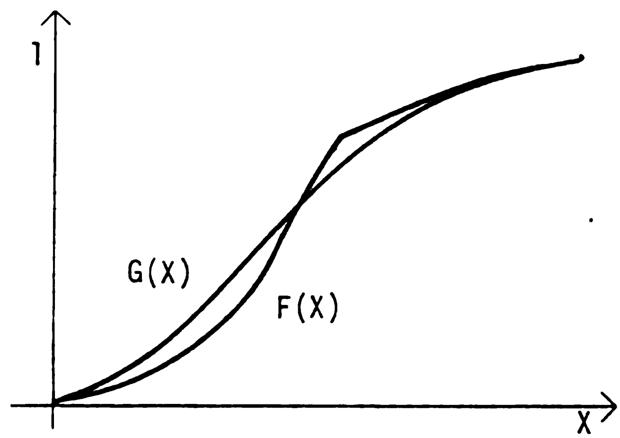



Figure 55.--Case where F dominates G under second order stochastic dominance.

From the statement of the necessary and sufficient conditions for second order stochastic dominance, it is clear that F and G can cross several times as long as the negative area between them (where F > G) to the left of any point remains smaller than the positive area (where G > F).

## A.8 Some Insights on Determination of Stochastic Dominance

a. From the condition $\int_a^Y S(X)dX > 0 \;\forall\; Y \in [a,b]$, it follows that the first nonzero value of $S(X) = G(X) - F(X)$ must be positive. Thus, if G(X) starts out less than F(X), then it is impossible for F(X) to dominate G(X) for all $U_i$ i=1 . . . , 3. This is one of the major weaknesses of risk theory.

b. If $M_F = M_G$, then $\sigma_F = \int X^2 dF(X) < \sigma_G = \int X^2 dG(X)$,[10] if F is to dominate G is the second degree.

c. If $M_F > M_G$, then $\sigma_F \gtrless \sigma$ if F dominates G.

## A.9 Third Order Stochastic Dominance

Arrow has argued[11] that decreasing absolute risk aversion is consistent with observed human behavior. Decreasing Absolute Risk Aversion (DARA) implies that the willingness to engage in a bet of fixed size increases with wealth. This is so because in the case of DARA, the odds demanded to enter into the bet diminish as wealth increases. Third order stochastic dominance can include this assumption about human behavior in the utility function and rank states of the world using the DARA assumption. As before, inclusion of this assumption allows the ranking of same states which previously could not be ordered.

For $U = U(X)$, absolute risk aversion is measured as

$$R_a = \frac{-U''(X)}{U'(X)}$$

and decreasing absolute risk aversion implies that $\partial R_a/\partial X < 0$. This is a weaker condition than $U'''(X) > 0$.[12] Thus, utility functions exhibiting DARA are a subset of $U_3$ (that is, the set of U where $U' > 0$, $U'' < 0$, $U''' > 0$). The conditions for determination of stochastic dominance are the same for $U_3$ and DARA when the means under F and G are the same.[13]

## A.10 Condition for Third Order Stochastic Dominance

F dominates G if and only if

a. $M_F \geq M_G$ and

b. $\int_a^e \int_a^Z S(Y)dYdZ \geq 0 \ \forall Z, \ e \in [a,b]$.

How this condition differs from that of second order stochastic dominance can be seen most clearly in Figure 56.

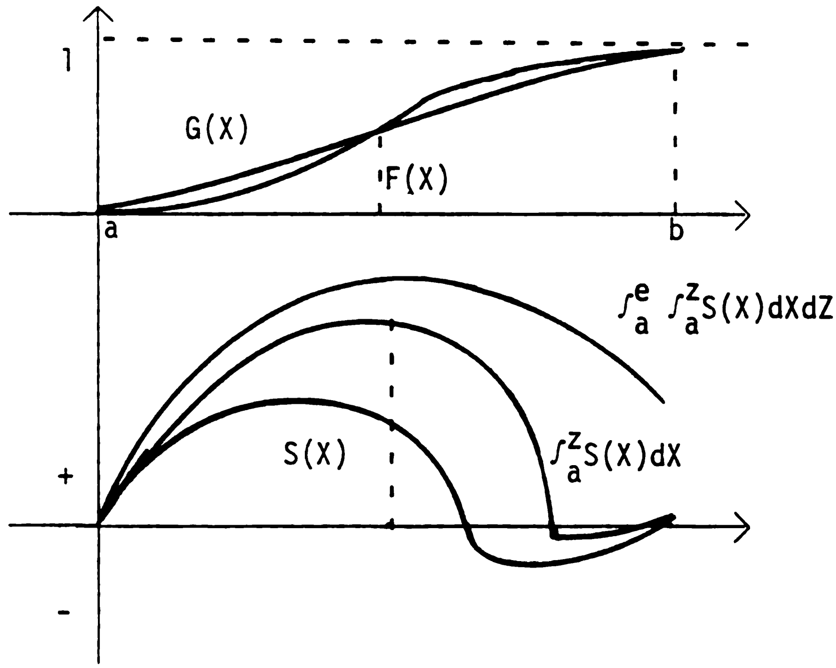



Figure 56.--Example of "rules" for determination of stochastic dominance.

From what has been presented above, it should now be clear that additional restrictions placed upon the utility function allow the set of distribution functions that can be ranked to be increased.

## A.11 Stochastic Dominance and the Level of Risk

Additional insights into the meaning of stochastic dominance and the times at which the existence of such dominance may be determined have been presented by Rothschild and Stiglitz.[14]

In the theory of expected utility maximization, a risk averter is defined as an agent having a concave utility function. If X and Y have the same mean but all risk averters prefer X to Y, for example, $EU(X) \geq EU(Y)$, then it is not unreasonable to say X is less risky than Y.

Under general conditions Rothschild and Stiglitz[15] have shown that the above definition of riskiness is equivalent to the two following formulations of increases in risk under general conditions developed below:

a. $EU(X) \geq EU(Y)$ if Y has more weight in the tails, that is, if the density function of Y is formed from the density function of X by taking some probability weight from the center of the density associated with X and shifting it to its tails to form the density for Y while leaving $E(Y) = E(X)$, then Y is more uncertain than X.

b. $EU(X) \geq EU(Y)$ if Y is equal to X plus noise, that is, $Y \overset{=}{_d} X + Z$, where "$\overset{=}{d}$" means "has the same distribution as" and Z is a random variable with the property that $E(Z/X) = 0$.

In showing the equivalence of the above definitions of greater risk, Rothschild and Stiglitz show that the partial ordering induced upon the distribution functions by the definition of an increase in risk in formation "a" is the same as the partial orderings induced by the other two definitions of an increase in risk.

Let g, a probability density function, be formed by shifting probability mass from the center of f, a probability density function, to its tails in such a way that the expected value of the random variable X is the same under both f and g. Let F and G be the cumulative distributions associated respectively with f and g.

Let

$$S(X) = g(X) - f(X) \quad \text{and} \quad S(X) = G(X) - F(X).$$

Then, under the condition that the random variable X has the same mean under g as f,

$$\int_a^b S(X) = \int_a^b (g(X) - f(X)) = \int_a^b X(g(X) - f(X)) = 0.$$

S(X) is called a mean preserving spread. If g is formed by a single mean preserving spread, then several properties hold. First, $S(a) = S(b) = 0$, where the interval [a,b] is the interval in which the random variable X has nonzero probability. Second, there is a Z such that for all X less than or equal to Z, $S(X) \geq 0$ and for all X greater than Z, $S(X) < 0$. Thus if g is formed from f by a single shift of probability mass toward its tails, the cumulative distributions associated with f and g will have the following shape:

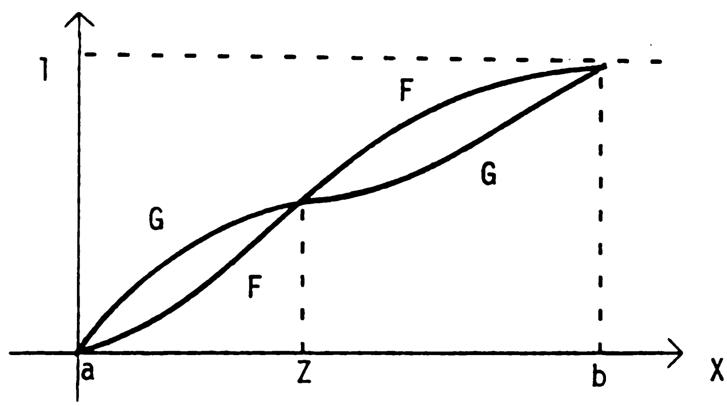



Figure 57.--Example of mean preserving spread.

Third, if $T(Y) = \int_a^Y S(X)dX$, $(Y \in [a,b])$, then $T(b) = 0$ since $T(b) = \int_a^b S(X)dX = X \int(X)\big|_a^b - \int_a^b XdS(X) = 0$. Points two and three together imply that $T(Y) \geq 0$, $a \leq Y < b$.

## A.12 The Integral Conditions

$T(b) = 0$ and $T(Y) \geq 0$, $a \leq Y < b$ are referred to as the integral conditions. It is shown by Rothschild and Stiglitz that if G is obtained from F by sequence of mean preserving spreads, then G - F satisfies the integral conditions. They next showed that if G - F satisfies the integral conditions, G could have been obtained from F by a sequence of mean preserving spreads. Thus, the integral conditions preserve the notion of transitivity.

Finally, the authors develop three partial orderings corresponding to the three concepts of increases in risk developed above, and proved their equivalence. It is shown that the orderings induced by the three definitions of increases in risk given above are each binary, reflexive, antisymmetric, and transitive. Each ordering is a partial ordering. Then it is shown that each pair of partial orderings is equivalent. These orderings are partial orderings over distributions with the same mean. Thus, if F and G have the same mean but T(Y) changes sign, then F and G cannot be ordered. That is, given two concave utility functions $U_1$ and $U_2$, there will be some risk-averse individuals from whom $U_1 \, dF(X) \leq U_1 \, dG(X)$ while at the same time there will be some risk-averse individuals for whom $U_2 \, dF(X) \geq U_2 \, dG(X)$.

The Rothschild and Stiglitz result may be extended by a generalization of Hanoch's Theorem.[16] The Rothschild formulation applies only to spreads which preserve the mean. It is not necessary that F(X) and G(X) have the same mean. It can be shown that if the mean of X is equal to or greater than the mean of Y and every risk averter prefers X to Y, that is, $EU(X) \geq EU(Y)$ for all concave U, then it is reasonable to assume that X is less risky than Y. This is analogous to the statement that Y has more weight in the tails than X if X and Y have density functions f and g, respectively.

Let f and g differ by a single spread s where $s = g - f$ and $S = G - F$ where $S(Z) = \int_a^b S(Z)dZ = G(Z) - F(Z)$. The properties of S are:

a. $S(a) = S(b) = 0$,

b. $S(X) \geq 0$ if $X \geq Z$
$S(X) < 0$ if $X > Z$.

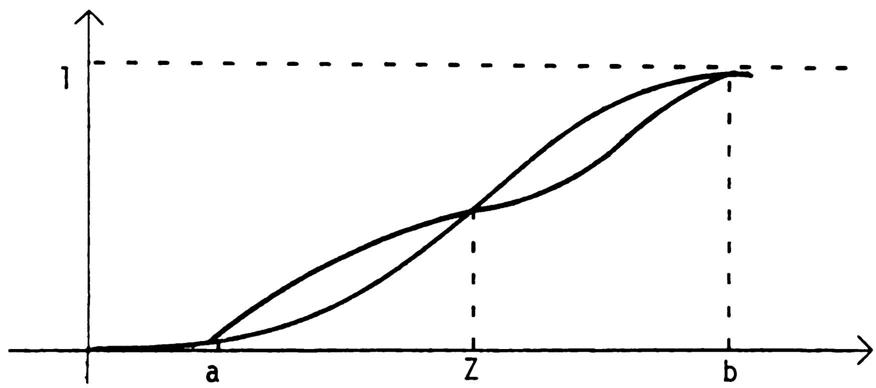



Figure 58.--Example of non-mean preserving spread.

If $T(Y) = \int_a^Y S(X)\, dX$, then

c. $T(b) > 0$ since

$T(b) = S(X)\, dX = XS(X) \mid_b^a - \int XdS(X) = 0 - \{E_G(X) - E_F(X)\} \geq 0.$

Conditions b and c together imply

d. $T(Y) > 0$.

Proof: $T(Y) = \int_a^Y S(X)\, dX$.

Let $Y \leq Z$ then from 'b' $T(Y) > 0$.

Let $Y > Z$ then $T(Y) = \int_a^Y S(X)\,dX + \int_Z^Y S(X)\,dX$.

Assume that for $Y > Z$ that $T(Y) < 0$ then $\int_a^Z S(X)\,dX +$

$$\int_Z^Y S(X)\,dX < 0 = \int_a^Z S(X)\,dX < \int_Z^Y S(X)\,dX.$$

Now $T(b) = \overset{Z}{\underset{b}{}}\int S(X)\,dX + \int_Z^b S(X)\,dX > 0$ (from c above) $=$

$$\int_a^Z S(X)\,dX > \int_Z^b - S(X)\,dX.$$

Thus $- \int_Z^b (X)\,dX < \int_a^Z S(X)\,dX < -\int_Z^Y S(X)\,dX \Rightarrow$

$$- \int_Z^b S(X)\,dX < - \int_Z^Y S(X)\,dX \quad \text{thus}$$

$$\int_Z^b S(X)\,dX > \int_Z^Y S(X)\,dX,$$

but $\int_Z^b S(X)\,dX < 0$ and is the smallest such number in the interval $(Z,b)$; thus $\int_Z^b S(X)\,dX > \int_Z^Y S(X)\,dX$ cannot hold and it must be that $T(Y) > 0$.

Let $T(b) \geq 0$ and $T(Y) \geq 0 \ \forall Y \in (a,b)$ be called the extended integral conditions.

Any cumulative distribution function G which can be obtained from F by use of the extended integral conditions will satisfy the above three definitions of increases in risk. In particular, if the extended integral conditions are met $E_F U(X) > E_G U(X)$.

Stiglitz,[17] in analyzing distributions which are close to one another, generalized the notation used to describe mean preserving spreads. Consider a family of distributions, $F(Y,r)$ where Y is a random variable. Let $F(Y,r_2)$ be derived from $F(Y,r_1)$ such that the two integral conditions hold, that is,

1. $\int_b^a [F(\theta,r_2) - F(\theta,r)]\, d\theta = 0$ and

2. $\int_0^Y [F(\theta,r_2) - F(\theta,r)]\, d\theta > 0 \;\forall Y \in (a,b)$.

In comparing distributions, which are close together, $F(\theta,r_2)$ represents a mean preserving increase in risk over $F(\theta,r_1)$ if

1. $\int_b^a F_r(\theta,r)\, d\theta = 0$

2. $\int_0^Y F_r(\theta,r)\, d\theta > 0 \;\forall Y \in (a,b)$, where F

where $F_r = \frac{\partial F(\theta,r)}{\partial r}\, dr$

This concept and notation will be used extensively.

## A.13 Extensions of the Concept of Increases in Risk

Jack Meyer,[18] in an attempt to extend the Rothschild and Stiglitz definition of increases in risk reformulated their definition of increases in risk as follows:

> G(X) is at least as risky as F(X) if F(X) is unanimously preferred or indifferent to G(X) by all agents who are at least as risk averse as a risk neutral agent . . . [where] increased risk aversion is taken to be increased risk aversion in the large. That is $U_1(X)$ is at least as risk averse in the large as $U_2(X)$ if $U_1(X)$ is a concave transformation of $U_2(X)$.[19]

Meyer believes that emphasis should be placed upon the relative level of risk aversion of the two decision makers since it is that relative level of aversion that determines whether F is preferred to G or vice versa. A new definition of risk aversion is presented by Meyer:

> Cumulative distribution G(X) is at least as risky as cumulative distribution F(X) if there exists some agent with a strictly increasing utility function U(X) (continuous and twice differentiable) such that for all agents more risk averse than he F(X) is preferred or indifferent to G(X).[20]

This definition of increasing risk does not require $U(X)$ to be concave and thus restrict attention to only those decision makers who are risk average in the Pratt-Arrow sense.[21] In developing the notion of increasing risk, the concept of a spread is introduced.

$S(X) = g(X) - f(X)$ is a spread if

a. $\int_a^b S(X)\, dX = 0$,

b. $S(X)$ changes sign at most twice in $(a,b)$,

c. $S(X) \geq 0$ for all $X \leq Z$ for some $Z$ contained in $(a,b)$, and

d. $S(X) > 0$ for some $X_0$ contained in $(a,Z)$.

Condition (a) states that the new density $g(X)$ will still satisfy the laws of probability for density functions. Conditions (b) through (d) guarantee that $g(X)$ will be formed from $f(X)$ by shifting weight to the tails of $f(X)$. In terms of cumulative distribution functions $G(X)$ and $F(X)$, their graphs could take the following general shape:

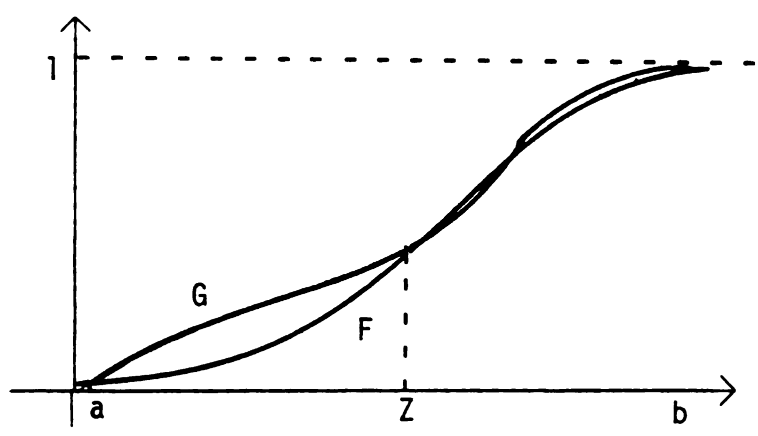



Figure 59.--Example of cumulative distributions under generalized spread.

Note that condition (b) does not require that G(X) cross over F(X) at $Z^1$. Also there is no restriction that requires the mean of X under G to be the same as that under F. Rothschild and Stiglitz require that $\int_a^b S(X)\,dX = \int_a^b XS(X)\,dX = 0$, which was required to show that T(b) = 0 (an integral condition). Meyer simply requires that $\int_a^b S(X)\,dX = 0$.

In developing his concept of increasing risk, Meyer proceeds along the following lines. Assume there exists an increasing, continuous twice differentiable r(X) such that

$$\int_0^Y \{G(X) - F(X)\}\,dr(X) \geq 0, \ \forall Y \in [a,b].$$

Then, if F is to dominate G it is necessary that

a. $\int_a^Y \{G(X) - F(X)\}\,dr(X) \geq 0, \ \forall Y \in [a,b].$

This is also sufficient.[22,23]

Let $Z = r(X)$, $(X = r'(Z))$ than (a) may be written as

$$\int_0^Y \{G(R^{-1}(Z)) - F(r^{-1}(Z)\}\,dr(r^{-1}(Z)) \geq 0 \ \forall Y \in [a,b] \text{ and}$$

b. $\int_0^Y \{B(r^{-1}(Z)) - Fr^{-1}(Z))\}\,dZ \geq 0 \ \forall Y \in [a,b].$

Let $G^*(Z) = G(r^{-1}(Z))$ and $F^*(Z) = F(r^{-1}(Z))$.

Then by direct substitution into (b) we get

$$\int_{r(a)}^{r(Y)} \{G^*(Z) - F^*(Z)Z\}\,dZ \geq 0 \quad \text{for all } Y \in [a,b].$$

Thus from Hadar Theorem 3[24] or Hanoch Theorem,[25] F* dominates G* for every U(Z) where U is nondeceasing and concave. In particular, from (b) we get b'. $\int_{r(a)}^{r(b)} \{G^*(Z) - F^*(Z)\}\, dU(Z) \geq 0$ $\forall$ concave, increasing twice differentiable U

b". $\int_a^b \{G(X) - F(X)\}\, dU(r(x)) \geq 0 = \int_a^b S(X)\, dU(X)$

where $U(X) = U(r(X))$.

Write $\frac{\partial U}{\partial X} = U' = \frac{\partial U(X)}{\partial r(X)} \frac{\partial (X)}{\partial X}$, and

then $\partial^2 \frac{U(X)}{\partial X^2} = U" = U"(r')^2 + U'r"$

then $\frac{U"}{U'} = \frac{U"r'}{U'} + \frac{r"}{r'}$

where $\frac{U"}{U'} = \frac{-U"}{U'} r' - \frac{r"}{r'}$ or $RU = RUr' + Rr$ where $RU \geq 0$

$Rr \geq 0_1 \quad r' > 0.$

Thus it is true that $R_U \geq R_r$, that is $U(X)$ represents all utility functions more risk averse than $r(X)$.

A second definition of increasing risk developed by Meyer is definition.[26]

> Cumulative distribution G(X) is at least as risky as cumulative distribution F(X) if G(X) can be obtained from F(X) by a finite sequence of cumulative distributions $F(X) = F_1(X)$, $F_2(X), \ldots, F(X) = F(X)$ such that each $F_1(X)$ differs from $F_{i-1}(X)$ by a single spread.

He continues by showing that $\int_a^Y \{G(X) - F(X)\}\, dr(X) \geq 0$ for all Y contained in (a,b) is the mathematical statement of both of his definitions of increasing risk if the probability functions of F(X) and G(X) cross a finite number of times. Finally, it is shown that[27]

> For any two cumulative distribution F(X) and G(X) such that their probability functions cross a finite number of times there exists an increasing continuous twice differentiable r(X) such that $\int_a^Y \{G(X) - F(X)\}\, dr(X) \geq 0 \; Y \in (a,b)$.

For any two cumulative distribution functions as defined above, either G is more risky than F or F is more risky than G. If F dominates G under this definition for a decision maker more risk averse than the person represented by $U_1$. This is not the same as saying that F is preferred to G by all risk-averse decision makers.

Meyer [28] extends the idea of choice among distributions and the idea of increases in risk and also relaxes the assumption of spreads, that is, the mechanical construction of G(X) from F(X) by some systematic rule. He starts by choosing U(X) such that

1. $r_1(X) \leq \frac{-U''(X)}{U'(X)} \leq r_\partial(X) \; \forall X$ .

If F(X) is preferred by all decision makers whose utility functions conform to (1), then it is true that for these decision makers

2. $\int_a^b (G(X) - F(X))\, U'(X)\, dX > 0$ .

If the inequality in (2) does not hold for some $U \equiv \frac{U_0'''}{U_0'} \in [r_1(X) r_2(X)]$, then it is not true that F is unanimously preferred by all decision makers whose utility functions conform with (1).

The problem of determining whether F is unanimously preferred to G by all decision makers whose utility functions conform to (1) can be reformulated as: Find that U where $r_1(X) \leq \frac{-U''(X)}{U'(X)} \leq r_2(X) \; \forall X$ and $U'(0) = 1$ and minimizes

3. $\int_a^b (G(X) - F(X))\, U'(X)\, dX$ .

If the minimum value of the integral is greater than or equal to zero, then F is unambiguously preferred to G.

Using Optimal Control Theory, Meyer develops the following rule for choice between G and F.[29] An optimal control $\frac{-U_0''}{U_0'}$ which maximizes $\int_0^1 (G(X) - F(X))\ U'(X)$ subject to $r_1(X) \leq [-U''(X)/U'(X)]\ Gr_2(X)$ and $U'(X) = 1$ is given by

$$\frac{-U_0''(X)}{U_0'(X)} = \begin{Bmatrix} r_1(X) \text{ if } \int X'[G(Y) - F(Y)]\ U'(Y)\ dY < 0 \\ r_2(X) \text{ if } \int X'[G(Y) - F(Y)]\ U(Y)\ dY \geq 0 \end{Bmatrix} .$$

If F and G cross a finite number of times, then the problem may be solved for the optimal $r = \frac{-U_0''(X)}{U_0'(X)}$. This U'(X) which minimized (3) is contained in either $r_1$ or $r_2$. Using the optimal U'(X) derived from the above theorem, the integral in (3) may be solved. If this solution is greater than zero, then F is unanimously preferred to G by all decision makers where $[r_1 \leq \frac{-U''(X)}{U'(X)} \leq r_2]$. If the solution to (3) is negative, this does not imply that G is unambiguously preferred to F. This technique provides a method for determining what class of decision makers unambiguously prefers F to G.[30]

## Footnotes--Appendix A

[1]J. Hadar and R. W. Russell, "Rules for Ordering Uncertain Prospects," American Economic Review 59,1 (1969): 25-34.

[2]C. Hanoch and H. Levy, "The Efficiency Analysis of Choices Involving Risk," Review of Economic Studies 36 (July 1969): 335-46.

[3]Actually, references prior to 1969 on stochastic dominance exist. For example: D. Blackwell and M. A. Girshick, Theory of Games and Statistical Decisions (New York: John Wiley and Sons, 1954).

[4]For example, see the pioneering efforts by Harry M. Markowitz, Portfolio Section: Efficient Diversification of Investment (New York: John Wiley and Sons, 1959).

[5]A. Sandmo, "On the Theory of the Competitive Firm Under Price Uncertainty," American Economic Review 61,1 (1971): 65-73; W. R. Russell and Paul E. Smith, "Taxation, Risk Taking, and Stochastic Dominance," Southern Economic Journal 36 (April 1970): 425-33.

[6]Ralph L. Keeney and Howard Raiffa, Decisions With Multiple Objectives: Preferences and Value Tradeoffs (New York: John Wiley and Sons, 1976).

[7]Hadar and Russell; Hanoch and Levy.

[8]G. Debreu, Theory of Value: An Axiomatic Analysis of Economic Equilibrium (New Haven: Yale University Press, 1959).

[9]Hanoch and Levy, p. 336.

[10]Ibid., p. 343.

[11]K. J. Arrow, Essays in the Theory of Risk Bearing (Chicago: Markham, 1971), especially chapter 3, "The Theory of Risk Aversion."

[12]This is so since $Ra(X) = \frac{-U''(X)}{U'(X)}$ and decreasing absolute risk aversion implies that $\frac{\partial Ra(X)}{\partial X} < 0$. But

$$\frac{\partial Ra(X)}{\partial X} = - \left\{ \frac{U'''(X)U'(X) - U''(X)U''(X)}{(U'(X))^2} \right\}.$$

This is negative whenever $U'''(X)U'(X) - U''(X)U''(X) > 0$ or $U'''(X) > \frac{U''(X)U''(X)}{U''(Z)}$. Thus for $\frac{\partial Ra(X)}{\partial X} < 0$, $U'''(X) > \frac{(U''(X))^2}{U'(X)}$, which obviously is a subset of all U such that $U'(X) > 0$, $U''(X) < 0$, $U'''(X) > 0$.

[13]G. A. Whitmore, "Third-Degree Stochastic Dominance," American Economic Review 60,3 (1970): 457-59.

[14]Michael Rothschild and Joseph E. Stiglitz, "Increasing Risk I: A Definition," Journal of Economic Theory 2 (September 1970): 224-34.

[15]Ibid.

[16]Hanoch, p. 343.

[17]Peter Diamond and J. Stiglitz, "Increase in Risk and Risk Aversion," Journal of Economic Theory 8,3 (1974): 337-60.

[18]Jack Meyer, "Increasing Risk," Journal of Economic Theory 11 (August 1975): 119-32.

[19]Ibid., p. 122.

[20]Ibid.

[21]J. Pratt, "Risk Aversion in the Small and Large," Econometrica 32 (1964): 122-36.

[22]Hadar, p. 30.

[23]Levy and Hanoch, p. 336.

[24]Hadar, p. 26.

[25]Hanoch and Levy, p. 338.

[26]Meyer, p. 124.

[27]Ibid., p. 128.

[28]Jack Meyer, "Choices Among Distribution," Journal of Economic Theory 14 (April 1977): 326-36.

[29]Ibid., p. 341.

[30](Withdrawn from body: Since the estimation of utility function is not precise, Meyer likens the development of a class of utility functions where F is preferred to G to the concept of confidence intervals in statistics.)

# BIBLIOGRAPHY

## BIBLIOGRAPHY

Abraham, Henry J. The Judicial Process. 2nd ed. New York: Oxford University Press, 1968.

Adelstein, Richard P. "Plea Bargain in Theory: A Behavior Model of the Negotiated Guilty Plea." Southern Economic Journal 6 (January 1978): 488-503.

Allen, R. G. D. Macro-Economic Theory: A Mathematical Treatment. New York: St. Martin's Press, 1968.

________, and Mishan, E. J. "Is the Substitution Term Ambiguous?" Economica 32 (May 1965): 215-22.

Alschuler, Albert W. "The Prosecutor's Role in Plea Bargaining." University of Chicago Law Review 36,50 (1968): 50-54.

American Bar Association, Advisory Committee on the Criminal Trial. Standards Relating to Pleas of Guilty. New York: American Bar Association, 1967.

Arrow, K. J. Essays in the Theory of Risk Bearing. Chicago: Markham, 1971.

Ashenfelter, Orley A., and Johnson, George E. "Bargaining Theory, Trade Unions, and Industrial Strike Activity." American Economic Review 59 (March 1969): 35-49.

Becker, Gary S. "Crime and Punishment: An Economic Approach." Journal of Political Economy 76 (March/April 1968): 169-217.

Bental, Benjamin, and Comay, Yochanan. "A Dynamic Bargaining Model." Paper presented at the Winter Meetings of the Econometric Society, October 1973. Mimeographed.

Bishop, Robert L. "Game Theoretic Analyses of Bargaining." Quarterly Journal of Economics 77 (November 1963): 559-602.

Blackwell, D., and Gershick, M. A. Theory of Games and Statistical Decisions. New York: John Wiley and Sons, 1954.

Blaine, Robert, and Schultz, David. Modern Mathematics and Economic Analysis. New York: W. W. Norton and Co., 1973.

Block, Michael K., and Lind, Robert C. "Crime and Punishment Reconsidered." The Journal of Legal Studies 4 (January 1975): 241-27.

Block, Michael K., and Lind, Robert C. "An Economic Analysis of Crimes Punishable by Imprisonment." The Journal of Legal Studies 4 (June 1975): 479-92.

Boulding, Kenneth. Conflict and Defense: A General Theory. New York: Harper, 1962.

Burger, Warren. "Address at the American Bar Association Annual Convention." New York Times, 11 August 1970, p. 1.

Cartter, Allan M. Theory of Wages and Employment. Homewood, Ill.: Irwin, 1959.

Chamberlain, Neil. Collective Bargaining. New York: McGraw-Hill Book Co., 1951.

Chiang, Alpha C. Fundamental Methods of Mathematical Economics. New York: McGraw-Hill, 1967.

Christensen, D. W., and Lau, L. J. "Transcendental Logarithmic Utility Functions." American Economic Review 65 (June 1975): 367-83.

Coase, R. H. "The Problem of Social Cost." Journal of Law and Economics 3 (October 1960): 1-44.

Coddington, Alan. "A Theory of the Bargaining Process: Comment." American Economic Review 56 (June 1966): 522-30.

Cole, George F. The American System of Criminal Justice. North Scituate: Duxbury Press, 1975.

Cross, John G. The Economics of Bargaining. New York: Basic Books, Inc., 1969.

________. "A Theory of the Bargaining Process." American Economic Review 55 (March 1965): 67-94.

________. "A Theory of the Bargaining Process: Reply." American Economic Review 56 (June 1966): 530-33.

Dahl, Robert. "The Concept of Power." Behavioral Science 2 (1957): 201-22.

Danziger, Sheldon, and Wheeler, David. "Malevolent Interdependence, Inequality, and Crime." M.I.T., September 1973. Mimeographed.

Dawson, Robert O. Sentencing: The Decision as to Type, Length, and Conditions of Sentence. Boston: Little, Brown and Co., 1969.

Debreu, G. The Theory of Value: An Axiomatic Analysis of Economic Equilibrium. New Haven: Yale University Press, 1959.

de Menil, George. Bargaining: Monopoly Power versus Union Power. Cambridge: M.I.T. Press, 1971.

Diamond, Peter, and Stiglitz, J. "Increases in Risk and Risk Aversion." Journal of Economic Theory 8 (July 1974): 337-60.

Ehrelick, Issac. "Participation in Illegitimate Activities: An Economic Analysis." Essays in the Economics of Crime and Punishment. New York: Columbia University Press, 1973.

Figlio, Robert M. "The Seriousness of Offenses: An Evaluation by Offenders and Nonoffenders." The Journal of Criminal Law and Criminology 66 (1975): 189-200.

Finney, D. J. Probit Analysis. 3rd ed. Cambridge: Cambridge University Press, 1971.

Fishburn, Peter C. "Convex Stochastic Dominance with Continuous Distribution Function." Journal of Economic Theory 7 (1974): 143-58.

________. "Separation Theorems and Expected Utilities." Journal of Economic Theory 11 (1975): 16-34.

Foldes, Lucien. "A Determinate Model of Bilateral Monopoly." Economica 31 (February 1964): 117-31.

Folk, H. W., and Wolfe, J. N. "The Ambiguity of the Substitution Term." Economica 31 (August 1964): 288-93.

Forst, Brian, and Brosi, Kathleen. "A Theoretical and Empirical Analysis of the Prosecutor." The Journal of Legal Studies 6 (January 1977): 177-92.

Fraser, D. A. S. Probability and Statistics. North Scituate: Dusbury Press, 1976.

Friedman, J. W. "On Experimental Research in Oligopoly." Review of Economic Studies 36 (October 1969): 399-416.

Frisch, Ragnar. Theory of Production. Chicago: Rand McNally and Co., 1965.

Goldfeld, Stephen, and Quandt, R. Nonlinear Methods in Econometrics. New Holland: North Holland Publishing Co., 1972.

Gould, John P. "The Economics of Legal Conflict." The Journal of Legal Studies 2 (June 1973): 279-300.

Green, Edward. Judicial Attitudes in Sentencing, A Study of the Factors Underlying the Sentencing Practice of the Criminal Court of Philadelphia. London: Macmillan and Co., Ltd., 1961.

Gumbel, E. J. "Ranges and Midranges." Annals of Mathematical Statistics 15 (1944): 414-22.

Gumbel, E. J., and Keeney, D. R. "The Extremal Quotient." Annals of Mathematical Statistics 21 (1950): 523-38.

Hadar, Joseph. Mathematical Theory of Economic Behavior. Reading: Addison Wesley Publishing Co., 1971.

Hadar, J., and Russell, W. R. "Applications in Economic Theory and Analysis." In Stochastic Dominance, pp. 293-334. Edited by G. A. Whitmore and M. C. Findlay. Lexington, Ky.: Lexington Books, 1978.

Hadar, Josef, and Russell, W. R. "Rules for Ordering Uncertain Prospects." American Economic Review 59 (March 1969): 25-34.

Hamermesh, Daniel S. "Who 'Wins' in Wage Bargaining?" Industrial and Labor Relations Review 26 (July 1973): 1146-49.

Hanoch, C., and Levy, H. "The Efficiency Analysis of Choices Involving Risk." Review of Economic Studies 36 (July 1969): 335-46.

Harsanyi, John C. "Approaches to the Bargaining Problem Before and After the Theory of Games: A Critical Discussion of Zeuthen's, Hicks', and Nash's Theories." Econometrica 24 (April 1956): 144-57.

Henderson, James M., and Quandt, Richard E. Microeconomic Theory: A Mathematical Approach. 2nd ed. New York: McGraw-Hill, 1971.

Hicks, John R. The Theory of Wages. 2nd ed. London: Macmillan Co., 1963.

Johnson, N., and Kotz, Samuel. Distributions in Statistics, Continuous Univariate Distributions. Vol. 2. New York: Houghton Mifflin Co., 1972.

Johnston, J. Econometric Methods. 2nd ed. New York: McGraw-Hill, 1972.

Joseph, M. L., and Willis, R. H. "An Experimental Analog to Two-Party Bargaining." Behavioral Science 8,2 (1963): 117-27.

Katzman, Martin. "The Economics of Defense Against Crime in the Streets." Land Economics 44 (November 1968): 431-40.

Keeney, Ralph L., and Raiffa, Howard. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: John Wiley and Sons, 1976.

Kihlstrom, R., and Mirman, Leonard. "Risk Aversion with Many Commodities." Journal of Economic Theory 8 (July 1974): 361-88.

Klonski, James R., and Mendelsohn, Allan. "The Allocation of Justice: A Political Approach." In The Politics of Local Justice, pp. 1-19. Edited by James R. Klonoski and Allan Mendelsohn. Boston: Little, Brown and Co., 1970.

Kmenta, Jan. Elements of Econometrics. New York: Macmillan Co., 1971.

Kogiku, K. C. Microeconomic Models. New York: Harper and Row, 1971.

Kraus, Alan, and Melnik, Arie. "A Sequential Decision Model of Bargaining Theory." Western Economic Journal 10 (December 1972): 359-75.

La Cava, Gerald J. "Improving the Mean-Variance Criterion Using Stochastic Dominance." Decision Sciences 7 (1976): 29-39.

Lachman, Judith A. "An Economic Model of Plea Bargaining in the Criminal Court System." Ph.D. dissertation, Michigan State University, 1975.

LaFave, Wayne. Arrest: The Decision to Take a Suspect into Custody. Boston: Little, Brown and Co., 1965.

Lancaster, Kevin. Consumer Demand, A New Approach. New York: Columbia University Press, 1971.

________. Mathematical Economics. New York: The Macmillan Co., 1968.

Landes, William M. "An Economic Analysis of the Courts." Journal of Law and Economics 14 (April 1971): 61-108.

________. "Rules for an Optimal Bail System: An Economic Approach." The Journal of Legal Studies 2,1 (1973): 79-106.

Laster, Clarence, Jr. "Pretrial Conference Division (Explanation of Precedure's [sic] and Forms)." Pretrial Conference Division, Wayne County, Michigan [post-1970]. Mimeographed.

Lipsey, R. G., and Lancaster, Kevin. "The General Theory of the Second Best." Review of Economic Studies 24 (October 1965): 11-32.

Luce, R. Duncan, and Raiffa, Howard. Games and Decisions. New York: John Wiley and Sons, 1957.

Mantel, Nathan. "Models for Complex Contingency Tables and Polychotomous Dosage Response Curves." Biometrics (March 1966): 83-95.

Markowitz, Harry M. Portfolio Section: Efficient Diversification of Investment. New York: John Wiley and Sons, 1959.

McFadden, D. "Conditional Logic Analysis of Qualitative Choice Behavior." In Frontiers of Econometrics. Edited by Paul Zarembka. New York: Academic Press, 1974.

McIntyre, Donald M., ed. Law Enforcement in the Metropolis, A Working Paper on the Criminal Law System in Detroit. Chicago: The American Bar Foundation, 1967.

Meyer, Jack. "Choices Among Distribution." Journal of Economic Theory 14 (April 1977): 326-36.

________. "Increasing Risk." Journal of Economic Theory 11 (August 1975): 119-32.

Michigan House Bill 5073, January 1, 1977. Section 227b 1, 2, 3.

Miller, Frank W. Prosecution: The Decision to Charge a Suspect With a Crime. The Report of the American Bar Foundation's Survey of the Administration of Criminal Justice in the United States. Boston: Little, Brown and Co., 1969.

Mills, James. "I Have Nothing to Do with Justice." Life, March 12, 1971, pp. 56-67.

Misra, P. N. "Relation Between Pearsonian Coefficients of Distributions of Least Squares Estimators and the Disturbance Term." Journal of the American Statistical Association 67 (September 1972): 662-63.

Mood, A.; Graybill, F.; and Boes, D. Introduction to the Theory of Statistics. 3rd ed. New York: McGraw-Hill, 1974.

Nagel, Stuart, with Marian Neef. "Some New Directions in Applying Economic Modeling to Judicial Research." Paper presented at the Roundtable on 'New Directions in Judicial Research,' Annual Meeting of the American Political Science Association, Chicago, Illinois, August 29-September 2, 1974.

Nash, John F., Jr. "The Bargaining Problem." Econometrica 18 (April 1950): 155-62.

Neuman, John W., and Morgenstern, Oskar. Theory of Games and Economic Behavior. New York: John Wiley and Sons, 1944.

Newman, Donald J. Conviction, The Determination of Guilt or Innocence Without Trial. Boston: Little, Brown and Co., 1966.

Patchen, Martin. "Models of Cooperation and Conflict: A Critical Review." Journal of Conflict Resolution 14 (September 1970): 389-407.

Pen, Jan. "A General Theory of Bargaining." American Economic Review 42 (March 1952): 24-42.

_______. The Wage Rate Under Collective Bargaining. Translated by T. S. Preston. Cambridge, Mass.: M. I. T. Press, 1966.

Peters, Edward. Strategy and Tactics in Labor Negotiations. New London: National Foremen's Institute, 1955.

Pindyck, R. S., and Rubinfeld, D. L. Econometric Modes and Economic Forecasts. New York: McGraw-Hill, 1978.

Plackett, R. L. "Linear Estimation from Censored Data." Annals of Mathematical Statistics 29 (1958): 131-42.

Pollak, Robert A. "Additive Von Neumann-Morgenstern Utility Function." Econometrica 35 (July-October 1967): 485-94.

Posner, Richard A. Economic Analysis of Law. Boston: Little, Brown and Co., 1972.

Pratt, J. "Risk Aversion in the Small and Large." Econometrica 32 (January 1964): 122-36.

Raiffa, Howard. Decision Analysis. Reading: Addison-Wesley, 1976.

Ramsey, J., and Gilbert, R. "A Monte Carlo Study of Some Small Sample Properties of Tests for Specification Error." Journal of the American Statistical Association 67 (March 1972): 180-86.

Remington, Frank J. Criminal Justice Administration. Indianapolis: Bobbs-Merrill Co., 1969.

Rhodes, W. "The Economics of Criminal Courts: A Theoretical and Empirical Investigation." The Journal of Legal Studies 5 (June 1976): 331-40.

________. "A Study of Sentencing in the Hennepin County and Ramsey County District Courts." The Journal of Legal Studies 6 (June 1977): 332-52.

Roberts, Blaine, and Schulze, David. Modern Mathematics and Economic Analysis. New York: W. W. Norton and Co., 1973.

Rothschild, Michael, and Stiglitz, Joseph E. "Increasing Risk I: A Definition." Journal of Economic Theory 2 (September 1970): 224-34.

Russell, William R., and Smith, Paul E. "Taxation, Risk Taking, and Stochastic Dominance." Southern Economic Review 36 (April 1970): 425-33.

Samuelson, Paul A. "Analytical Aspects of Anti-Inflation Policy." The American Economic Review 50 (May 1960): 177-94.

Sandmo, A. "On the Theory of the Competitive Firm Under Price Uncertainty." American Economic Review 61 (March 1971): 65-73.

Saraydar, Edward. "A Certainty-Equivalent Model of Bargaining." Journal of Conflict Resolution 15 (September 1971): 281-97.

________. "Uncertainty, Bargaining, and the Strategy of Commitment." Zeitschrift fur Nationalokonomie 31 (1971): 309-22. (Also University of Western Ontario Department of Economics Reprint No. 95.)

Schelling, Thomas C. The Strategy of Conflict. Cambridge, Mass.: Harvard University Press, 1960.

Shubik, Martin, Ed. Game Theory and Related Approaches to Social Behavior. New York: John Wiley and Sons, 1964.

State Journal [Lansing, Michigan]. January 5, 1979, Sec. 1.

Stiglitz, Joseph E. "Behavior Toward Risk with Many Commodities." Econometrica 37 (October 1969): 660-67.

Task Force Reports: The Courts. Task Force on Administration of Justice--The President's Commission on Law Enforcement and Administration of Justice. Washington, D.C.: Government Printing Office, 1967.

Theil, H. "Estimates and Their Sampling Variance of Parameters of Certain Heteroschedastic Distribution." Review de L'Institut International de Statistique 19 (1951): 141-47.

________. "A Multinomial Extension of the Linear Logit Model." International Economic Review 10 (1969): 251-59.

________. Principles of Economics. New York: John Wiley and Sons, 1971.

Thurow, Lester C. "Equity Versus Efficiency in Law Enforcement." Public Policy 18 (Summer 1970): 451-62.

Tobin, James. "Liquidity Preference as Behavior Towards Risk." Review of Economic Studies 67 (February 1958): 65-86. Also in Risk Aversion and Portfolio Choice. Edited by Donald J. Hester and James Tobin. New Haven: Yale University Press, 1967.

Trebach, Arnold S. The Rationing of Justice. New Brunswick, N.J.: Rutgers University Press, 1964.

Varien, Hal R. Microeconomics Analysis. New York: W. W. Norton & Co., 1978.

Vetri, Dominick. "Guilty Pleas Bargaining: Compromises by Prosecutors to Secure Guilty Pleas." University of Pennsylvania Law Review 112 (1964): 865-68.

Vickrey, William S. Microstatistics. New York: Harcourt, Brace and World, Inc., 1964.

Vickson, R. G. "Stochastic Dominance Tests for Decreasing Absolute Risk-Aversion II: General Random Variables." Management Science 23 (January 1977): 478-89.

Virtue, Maxine. Survey of Metropolitan Courts, Final Report. Ann Arbor: University of Michigan Press, 1962.

Whitmore, G. A. "Third Degree Stochastic Dominance." American Economic Review 60,3 (1920): 457-59.

Whitmore, G. A., and Findley, M. C., ed. Stochastic Dominance. Lexington Books, 1978.

Yamane, Taro. Mathematics for Economists. 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Zeisel, Hans. "Methodological Problems in Studies of Sentencing." Law and Society Review 3 (May 1969): 621-31.

Zeuthen, Frederik. Problems of Monopoly and Economic Warfare. London: George Routledge and Sons, 1930.