University of Miami Law Review

Volume 31 | Number 5 Article 5

11-1-1977

Decision Theory and the Pre-Trial Release Decision in Criminal Cases

Stuart Nagel

Marian Neef

Sarah Slavin Schramm

DECISION THEORY AND THE PRE-TRIAL RELEASE DECISION IN CRIMINAL CASES*

STUART NAGEL,** MARIAN NEEF,*** AND SARAH SLAVIN SCHRAMM****

This article sets a framework for analyzing how a rational judge sets bond in individual criminal cases in order to maximize given goals. The assumed goals relate to maximizing the probabilities of the defendant appearing in court and his being released prior to trial and also minimizing the costs of rearresting bail-jumpers, pretrial crime-committing, jail maintenance, lost gross national product, and the bitterness that comes from being held in jail prior to trial when the case is dismissed or the defendant acquitted. The goal maximization is analyzed in the context of elementary decision theory working with both non-monetary and monetary values. The article is also concerned with deducing a set of policy recommendations that are designed to cause arraignment judges to become more sensitive to avoiding errors of holding defendants who would appear, relative to avoiding errors of releasing defendants who would fail to appear. Furthermore, the authors suggest that the concepts and methods of decision theory which are applied to the pre-trial release decision can also be used by analogy to analyze other criminal justice decisions that involve contingent events; for example, those decisions made by a

---

* This is the third article in a series of three dealing with the application of decision theory to legal process decision-making. The first dealt with group decision-making. *See* Nagel & Neef, *Deductive Modeling to Determine an Optimum Jury Size and Fraction Required to Convict*, 1975 WASH. U.L.Q. 979 (1976). The second article dealt with two-person decision-making. *See* Nagel & Neef, *Plea Bargaining, Decision Theory, and Equilibrium Models*, 51 IND. L.J. 987, 52 IND. L.J. 1 (1976). This present article deals with one-person decision-making, which is the perspective most broadly applicable to the legal process since so many important discretionary decisions are made by a single police officer, lawyer, judge, or individual members of a group.

Thanks are owed to the LEAA National Institute for Law Enforcement and Criminal Justice, the Ford Foundation Public Policy Committee, and the University of Illinois Law and Society Program for financing various aspects of the legal policy research of which this paper is a part, although none of them are responsible for the ideas advocated here. Thanks are also owed for comments on earlier drafts to Leslie Wilkins of SUNY-Albany, Allan Goldman of the National Bureau of Standards, Herbert Miller and William McDonald of the Georgetown Law School, Paul Lermack of Bradley University, Daniel Fried of the Yale Law School, and Joseph Ebersole of the Federal Judicial Center.

** Professor of Political Science at the University of Illinois and member of the Illinois bar.

*** Ph.D. candidate in political science at the University of Illinois.

**** Ph.D. candidate in political science at George Washington University.

police officer to arrest rather than issue a summons, by a prosecutor or defense counsel to go to trial rather than accept an out-of-court settlement, and by a sentencing judge or parole board to incarcerate or continue incarceration.

I. INTRODUCTION 1434
II. THE RELEASING AND BOND-SETTING DECISION IN INDIVIDUAL CASES 1436
A. *The Release or Hold Decision* 1437
1. NON-MONETARY VALUES 1437
2. APPLICATIONS AND VARIATIONS 1441
3. MONETARY VALUES 1448
B. *The Bond-Setting Decision* 1459
C. *Increasing the Expected Value of Releasing in Individual Cases* 1463
1. THE GENERAL PERSPECTIVE 1463
2. RAISE AND CLARIFY THE PROBABILITY OF APPEARANCE 1467
3. MAKE MORE VISIBLE THE ERRORS AND COSTS OF HOLDING DEFENDANTS WHO WOULD APPEAR 1469
4. DECREASE THE COSTS OF RELEASING DEFENDANTS WHO FAIL TO APPEAR 1472
III. THE BOND-SETTING DECISION ACROSS CASES 1474
A. *Non-Discretionary Bond Schedules* 1474
1. THE PROBLEM 1474
2. THE GOAL AND A SOLUTION 1476
3. VARIATIONS ON THE SOLUTION 1479
B. *The Optimum Percentage to Release* 1482
IV. SOME TENTATIVE CONCLUSIONS 1484
V. APPENDICES 1487
A. *Glossary of Symbols* 1487
B. *Basic Formulas Used* 1490

I. INTRODUCTION

In recent years, much concern has been expressed with regard to the lack of uniformity and effectiveness of sentencing in criminal cases. As both a cause and a result of that concern, there has been an increase in the number of studies of the sentencing decision especially (1) studies seeking to describe, explain, and decrease the non-uniformity and (2) studies seeking to determine the effects of diverse sentences, including benefit-cost analyses of would-be criminal behavior in light of the deterrent effect of possible conviction and sentencing.[1] The sentencing stage of the criminal justice pro-

1. On disuniformity in sentencing, see M. FRANKEL, CRIMINAL SENTENCES: LAW WITHOUT ORDER (1973). On the lack of effectiveness of sentences designed to rehabilitate, see Martinson, *What Works? Questions and Answers About Prison Reform*, 35 THE PUB. INTEREST 22 (1974). On the possible effectiveness of sentences designed to deter, see F. ZIMRING & G. HAWKINS, DETERRENCE (1973).

cess, however, may be a marvel of uniformity and effectiveness in comparison to the pre-trial release stage. In spite of that, there has been little analysis of the great diversity from judge to judge with regard to pre-trial release. There has likewise been little analysis of the effects of bond-setting on the probability of a defendant appearing in court. For awhile, though a flurry of studies did occur dealing with release on recognizance and preventive detention.[2] A major purpose of this article is to analyze the pre-trial release decision from a decision theory perspective which emphasizes the possible benefits and costs to arraignment judges and to society under various circumstances.

More specifically, decision theory can be defined as the study of which of various available decisions should be reached in order to maximize benefits minus costs in light of the probable occurrence of uncertain events.[3] In the context of pre-trial release, the available decisions are: (1) to release on a low bond or on the defendant's own recognizance; or (2) to hold on a high bond or on a non-bondable charge. The key probabilistic events are whether or not the defendant will appear in court, and whether or not the defendant will commit a crime while released. Although this article is primarily concerned with the pre-trial release decision, much of the analysis is applicable by analogy to other stages in the criminal justice process which also involve probabilistic decisions. Those stages include a police officer's decision to make an arrest contingent on the approval of his superior officers, a prosecutor's decision to plea bargain contingent on the probability of obtaining a conviction, a juror's

---

2. Ares, Rankin, & Sturz, *The Manhattan Bail Project: An Interim Report on the Use of Pre-Trial Parole,* 38 N.Y.U.L. REV. 67 (1963); *Preventive Detention: An Empirical Analysis,* 6 HARV. CIV. RIGHTS, CIV. LIB. L. REV. 289 (1971); Wald, *The Right to Bail Revisited: A Decade of Promise Without Fulfillment,* in THE RIGHTS OF THE ACCUSED 177-205 (Nagel ed. 1972).

3. For general discussions of decision theory, see S. RICHMOND, OPERATIONS RESEARCH FOR MANAGEMENT DECISIONS 527-60 (1968); W. BAUMOL, ECONOMIC THEORY AND OPERATIONS RESEARCH 550-68 (1965); H. RAIFFA, DECISION ANALYSIS: INTRODUCTORY LECTURES ON CHOICE UNDER UNCERTAINTY (1968); R. MACK, PLANNING ON UNCERTAINTY: DECISION MAKING IN BUSINESS AND GOVERNMENT ADMINISTRATION (1971). For discussions of decision theory applied to pre-trial release and other criminal justice matters, see Landes, *The Bail System: An Economic Approach,* 2 J. OF LEGAL STUD. 79 (1973); J. Locke, Compilation and use of Criminal Court Data in Relation to Pre-Trial Release of Defendants: Pilot Study 112-15 (1970); G. MONKMAN, READINGS IN CORRECTIONAL ECONOMICS (1973); Fried, et al., *Jury Selection: An Analysis of Voir Dire,* in THE JURY SYSTEM: A CRITICAL ANALYSIS (Simon ed. 1975); and Stover and Brown, *Reducing Rule Violations by Police, Judges, and Corrective Officials* in CRIMINAL JUSTICE MODELING (Nagel ed. 1977).

decision to convict contingent on the probability that the defendant is guilty, and a parole board's decision to release contingent on the inmate's likelihood of repeating his crime especially while he is on parole.

In addition to providing a better understanding of decision theory and the pre-trial release decision, the analysis presented in this paper is designed to serve a number of functions relevant to improving the pre-trial release process. Through decision theory as the basis for gathering data from arraignment judges, a researcher can indicate to the judges various types of biases they might have in their pre-trial release decisions but not be clearly aware of, especially with regard to their relative concern for avoiding an error of holding someone who would appear versus releasing someone who would have failed to appear. The decision theory perspective also provides a means for determining the implicit threshold probabilities which various judges have in making pre-trial release decisions. Revealing those probabilities to a set of judges on the same court may help to bring the more deviant judges closer to the average. Knowing those probabilities may also be a step to explaining the causes of the variation among judges, cases, and places. That kind of causal analysis can aid in determining the effects on decisions of changes in benefits, costs, or probabilities or of changes in their visibility. The benefit-cost aspects may also clarify the need for more pre-trial release of marginal defendants. The complete analysis may be especially useful in developing objective bond schedules or charts analogous to the objective decision-making charts used in workmen's compensation cases and increasingly being proposed for flat sentencing in criminal cases. Such bond schedules could conceivably pinpoint the bonds that should be set for various crimes in order to maximize the benefits of defendants appearing in court minus the costs of having to hold defendants in jail.

II. THE RELEASING AND BOND-SETTING DECISION IN INDIVIDUAL CASES

The bond-setting decision is an especially good decision making situation to illustrate various aspects of decision theory because it involves many contingent events, many alternative decisions rather than a simple dichotomy, both monetary and non-monetary values, both individualized cases and different case types, and both descriptive and optimizing elements.[4] These features of the bond-setting decision will be defined and clarified in the following article. We begin with the individual case where judicial discretion is important, and then later deal with non-discretionary bond schedules where legislatures or state supreme courts specify fixed bonds for various case types. Within the individual case, we first treat the release-versus-hold decision from both a non-monetary and a monetary perspective, and we then deal with the bond-setting decision which is more complicated than the simple dichotomy of releasing or holding.

A. *The Release or Hold Decision*

1. NON-MONETARY VALUES

The decision to release a defendant can take the form of releasing the defendant on his own recognizance without any bond (ROR) or setting a bond low relative to the defendant's ability to pay. The decision to hold can take the form of labeling the case a "no bond allowed" case or setting a bond high relative to the defendant's ability to pay. Figure 1 gives the payoff matrices for two hypothetical judges in the same case. A payoff matrix shows the satisfaction or dissatisfaction received or perceived by a decision-maker or a collectivity from each available decision and each possible occurrence of some uncertain event. A payoff matrix is a useful way of analyzing decisions, not necessarily a way of explicitly making them. In the pre-trial release context, there are two alternative decisions available, namely release or hold. There are likewise two alternative categories on the contingent event: (1) the defendant would appear if released; or (2) the defendant would fail to appear if released. The cells indicate the relative satisfaction or dissatisfaction received by each judge if he releases the defendant who then fails

4. Books on pre-trial release, bond setting, and related matters include STUDIES ON BAIL (Foote ed. 1966), especially at 4-15 and 74-88; P. WICE, FREEDOM FOR SALE: A NATIONAL STUDY OF PRE-TRIAL RELEASE (1974), especially at 25-34; D. FREED & P. WALD, BAIL IN THE UNITED STATES (1964); ABA, STANDARDS RELATING TO PRE-TRIAL RELEASE (1968); R. GOLDFARB, RANSOM: A CRITIQUE OF THE BAIL SYSTEM (1965); K. BOTTOMLEY, PRISON BEFORE TRIAL: A STUDY OF REMAND DECISIONS IN MAGISTRATE COURTS (1970); W. THOMAS, A DECADE OF BAIL REFORM (1976); B. MAHONEY, AN EVALUATION OF POLICY RELATED RESEARCH ON THE EFFECTIVENESS OF PRETRIAL RELEASE PROGRAMS (1975); J. MULLEN, PRE-TRIAL SERVICES: AN EVALUATION OF POLICY RELATED RESEARCH (1975); Silverstein, *Bail in the States: A Field Study and Report*, 50 MINN. L. REV. 621 (1966).

to appear (cell a); if he releases the defendant who does appear (cell b); if he holds the defendant when he would have failed to appear if released (cell c); and if he holds the defendant when he would have appeared if released (cell d). The most satisfying occurrence is anchored at +100, and the most dissatisfying is anchored at —100. The cell entries shown in Figure 1 are hypothetical, but Figure 2 to be discussed later deals with how such values may be empirically derived.

FIGURE 1. DECISION THEORY PAYOFF MATRICES AS PERCEIVED BY TWO ARRAIGNMENT JUDGES DECIDING WHETHER OR NOT TO RELEASE A DEFENDANT

1A. A JUDGE WHO IS MORE WORRIED ABOUT HOLDING A GOOD-RISK DEFENDANT THAN RELEASING A BAD-RISK DEFENDANT (oriented toward avoiding type 1 errors)

		PROBABILITY OF APPEARANCE (PA)		EXPECTED VALUE IF PA = .6
		Would Fail to Appear	Would Appear	
ALTERNATIVE DECISIONS AVAILABLE	Release via ROR or Low Bond	a —50	b +100	(.4) (—50) + (.6) (+100) = +40
	Hold via No or High Bond	c +75	d —100	(.4) (+75) + (.6) (—100) = —30

1B. A JUDGE WHO IS MORE WORRIED ABOUT RELEASING A BAD-RISK DEFENDANT THAN HOLDING A GOOD-RISK DEFENDANT (oriented toward avoiding type 2 errors)

		PROBABILITY OF APPEARANCE (PA)		EXPECTED VALUE IF PA = .6
		Would Fail to Appear	Would Appear	
ALTERNATIVE DECISIONS AVAILABLE	Release via ROR or Low Bond	a —100	b +25	(.4) (—100) + (.6) (+25) = —25
	Hold via No or High Bond	c +100	d —10	(.4) (+100) + (.6) (—10) = +34

Cells indicate relative satisfaction of each occurrence with the most satisfying anchored at +100 and the most dissatisfying anchored at —100.

The judge in Figure 1A is more worried about holding a good-risk defendant than releasing a bad-risk defendant, as indicated by the fact that he gets the most dissatisfaction from cell d. On the other hand, the judge in Figure 1B is more worried about releasing a bad-risk defendant than holding a good-risk defendant,[5] as indicated by the fact that he gets the most dissatisfaction from cell a. We assume that both judges are hearing the same case in the same city so that the differences in their perceived payoff values reflect their attitudinal differences. Otherwise, the differences might reflect the severity of the defendant's criminal behavior since the same judge could have a payoff matrix like 1B for a homicidal maniac, but a payoff matrix like 1A for a jaywalker. Likewise, a judge in a city that has high holding costs relative to releasing costs might have a payoff matrix like 1A, but a judge in a city that has high releasing costs relative to holding costs might have a payoff matrix like 1B. Holding costs in this context might refer to jail upkeep, lost earnings, and bitterness due to mis-arrests, whereas releasing costs refer to the costs due to rearresting no-shows and the monetary and psychological costs of crime committed by released defendants.

Suppose both judges perceive the defendant as having a probability of appearing (or PA) of about .60. If either judge were to be confronted with ten such defendants, this means about six would appear for their trial date and four would fail to appear. If the same defendant were to be given ten opportunities, this means about six times he would appear and four times he would fail to appear. Thus, the expected values for Judge 1A of releasing ten such defendants would be: four times he would suffer a —50 dissatisfaction; six times he would receive a +100 satisfaction; and he would thus average a +40 expected value from releasing our hypothetical defendant. Likewise, the expected values for Judge 1A of holding ten such defendants would be: four times he would receive a +75 satisfaction; six times he would suffer a —100 dissatisfaction; and he would

5. Judge 1A is more oriented toward avoiding a type 1 error than a type 2 error, whereas Judge 2A is more oriented toward avoiding a type 2 error than a type 1 error. A type 1 error involves rejecting a true hypothesis, whereas a type 2 error involves accepting a false hypothesis. The basic criminal justice system hypothesis or presumption is that the defendant is innocent and that he will appear in court. Rejecting that hypothesis when it is true (a type 1 error) means holding a defendant who would have appeared. Accepting that hypothesis when it is false (a type 2 error) means releasing a defendant who should have been held.

thus average a —30 expected value from holding our hypothetical defendant, assuming in both the releasing and holding situation that he could be made aware of the consequences of his actions. The same kind of expected value calculations could be done with Judge 1B. In general, an expected value is the benefits or costs associated with an action or decision discounted by or multiplied by the probability that the benefits or costs will occur.

Given the logical assumption that any judge or any person will prefer to choose the action or alternative decision that gives him the highest expected value, Judge 1A will logically prefer to release the hypothetical defendant with a probability of appearing of .60, and Judge 1B will prefer to hold such a hypothetical defendant. Rather then ask whether a given judge will release or hold a given defendant, the more interesting question is what is the threshold probability of appearance (PA*) that has to be met before Judge 1A or 1B will release a defendant. To calculate PA* for either judge, all we have to do is solve for PA in the equation (1—PA)(a) + (PA)(b) = (1—PA)(c) + (PA)(d) since at that PA level, the expected value of releasing exactly equals the expected value of holding.[6] Thus, for Judge 1A, his PA* or threshold probability of appearance equals .385; whereas for Judge 1B, his PA* equals .851. This means Judge 1A will release (or should release if he wants to maximize his expected values) any defendant who has a .39 or higher chance of appearing, and will hold any defendant who has a .38 or lower chance of appearing. On the other hand, Judge 1B will release (or should release if he wants to maximize his expected values) any defendant who has a .86 or higher chance of appearing, and will hold

6. If one solves for PA in the equation which equalizes the expected value of releasing and the expected value of holding, then the solution is PA* = (a—c)/(a—b—c+d). The symbol PA is shown with a star to indicate this is the value of the probability of appearing when the expected values of releasing and holding are equal. The proof of this formula is as follows:

1. (1—PA)(a) + (PA)(b) = (1—PA)(c) + (PA)(d)
2. a—Pa + Pb = c—Pa + Pd
   (Removing the parentheses, and substituting P for PA)
3. —Pa+Pb+Pc—Pd = c—a
   (Transposing)
4. P(—a+b+c—d) = c—a
   (Factoring out P)
5. P = (c—a)/(—a+b+c—d)
   (Dividing both sides by a+b+c—d)
6. PA* = (a—c)/(a—b—c+d)
   (Multiplying numerator and denominator of right side by —1)

any defendant who has a .85 or lower chance of appearing. Judge 1A probably releases a substantially higher percentage of the defendants who appear before him than Judge 1B does if their judicial behavior reflects their differential values and they face roughly the same defendants.

2. APPLICATIONS AND VARIATIONS

Actual judges could be positioned with regard to their orientation toward avoiding type 1 errors versus avoiding type 2 errors by calculating for each judge what his threshold or equilibrium PA* is. The higher his PA* is, the more he is oriented toward holding defendants. To obtain the values of judges for insertion into matrices like those shown in Figure 1 would involve asking them questions like those shown in Figure 2. The questions are in a form that facilitates mailed responses although they also could be administered in person. Asking judges directly as to what probability of appearance they require in order to release a defendant is likely to yield less reliable responses than this more indirect approach, although both approaches can be used.[7] The questionnaire can also contain other questions relating to the bond setting process including hypothetical bond-setting situations as described below in Section 1B.

Instead of having both an upper anchor at +100 and a lower anchor at —100, one might use just an upper anchor at +100 or just a lower anchor at —100. Doing so allows more freedom at the other end of the scale rather than sometimes artificially saying that the worst payoff has the same value as the best payoff but is opposite in sign. The worst payoff, however, is the same as the best payoff but opposite in sign if a judge says the cost of a type 1 error (of holding a defendant who would show up) is the holding costs incurred minus the releasing costs saved, and the cost of a type 2 error (of releasing a defendant who fails to show) is the releasing costs incurred minus the holding costs saved. It would not generally be meaningful to assign the worst payoff a value of zero because doing so might result in the next to the worst payoff having a positive

7. On various methodologies for obtaining payoff matrix data, see Huber, *Methods for Quantifying Subjective Probabilities and Multi-Attribute Utilities*, 5 DECISION SCIENCES 430 (1974); and P. KOTLER, MARKETING DECISION MAKING: A MODEL BUILDING APPROACH 583-95 (1971).

FIGURE 2. SOME QUESTIONS FOR OBTAINING A JUDGE'S VALUES REGARDING PRETRIAL RELEASE DECISIONS

General Statement: There are two things that could happen if you *release* someone. One is that he will appear in court when he is supposed to (possibility a); the other is that he will fail to appear in court when he is supposed to (possibility b). There are likewise two things that could have happened when you *hold* someone in jail. One is that he would have appeared in court if he had been released (possibility c); the other is that he would have failed to appear if he had been released (possibility d).

Question	Response
1. Which of these four possibilities would give you the most satisfaction? Let's score this one +100 to have an upper anchor point.	(a, b, c, or d)
2. Which of these four possibilities would give you the most *dis*satisfaction? Let's score this one —100 to have a lower anchor point.	(a, b, c, or d)
3. Which of these four possibilities would give you the next to the most satisfaction?	(a, b, c, or d)
4. In view of the satisfaction you receive from the most satisfying occurrence, how would you rate this next most satisfying occurrence on a scale from 0 to +100, where 0 means neutral, and +100 means most satisfying?	(a number from 0 to +100)
5. Which of these four possibilities would give you the next to the most dissatisfaction?	(a, b, c, or d)
6. In view of the dissatisfaction you receive from the most dissatisfying occurrence, how would you rate this next most dissatisfying occurrence on a scale from 0 to —100, where 0 means neutral, and —100 means most dissatisfying?	(a number from 0 to —100)
7. Are you sure you want your next to the best choice to be that close or that far away from the best choice, and that close or that far away from the neutral position? If not, please revise your answer accordingly.	(possible revised number for question 4)
8. Are you also sure you want your next to the worst choice to be as close or as far away from your worst choice as you initially put it? If not, please revise that evaluation accordingly.	(possible revised number for question 6)

value even though it is an undesirable payoff in an absolute as well as a relative sense.[8]

8. As a variation on Figure 2, instead of beginning by asking which of the four alternatives would give the most satisfaction, the questions can begin by saying "Which of these four

As an alternative to the set of questions shown in Figure 2 which some respondents may find difficult to handle because of the abstract nature of the numbers and the sometimes undesirable tendency to make the worst payoff the same as the best payoff but opposite in sign, one can talk in terms of dollars willing to be paid for each result. Thus, each judge would be asked, how much money would you be willing to pay out of your own pocket in order to avoid whatever hurt you might feel if a defendant you release fails to appear in court (possibility a)? Likewise, how much would you be willing to pay in order to be assured that a defendant you release will appear in court (possibility b)? How much in order to be assured that the defendant you hold would have failed to appear in court (possibility c)? Finally, how much would you be willing to pay in order to avoid the hurt of knowing that the defendant you hold would have appeared in court if he had been released (possibility

alternatives would give you some satisfaction, and which would give you some dissatisfaction?" Then the questions can ask, "Of the satisfying alternatives, which is the most satisfying?" the questionnaire could then determine a numerical value for the second most satisfying as Figure 2 does. Then one would ask, "Of the dissatisfying alternatives, which is the most dissatisfying?", and then likewise determine a numerical value for the second most dissatisfying. That procedure in effect partitions the decisions the respondent has to make into smaller more manageable decisions of determining direction, rank within direction, and then relative numerical value for the lesser rank, rather than trying to determine the most satisfying alternative first which is a question that combines both direction and rank. The respondent might also be more comfortable discussing the dissatisfying alternatives before the satisfying since people may tend to think more in terms of avoiding relatively bad errors in making decisions than in trying to maximize good results. In terms of Figure 1, these questions involve determining: (1) where to put a minus and where to put a plus; (2) which minus should be a double minus, and which plus should be a double plus; (3) how the single minus should be scored if the double minus is scored —100; and (4) how the single plus should be scored if the double plus is scored +100.

Another approach that has been used to assign payoff values to the alternative possibilities stemming from a decision theory problem is: (1) rank the payoffs from the most desirable to the least desirable; (2) determine the relative distances between each payoff; (3) assign a value of 100 to the most desirable and a value of 0 to the least desirable; (4) determine which one of the payoff outcomes has close to an indifferent value, meaning it produces neither satisfaction nor dissatisfaction; (5) subtract the numerical value of that indifferent payoff from each of the other numerical values, giving positive values to payoffs above the indifferent payoff and negative values to those below. This method is used by R. Tanter, Evaluation and Anticipation of Choice in International Crisis Management (1975) (unpublished paper available from the author at the University of Michigan Political Science Department). The results, however, become distorted if step 4 does not involve a truly indifferent payoff. Step 2 is difficult to execute if more than a pair of payoffs are being compared at once. Tanter also asks respondents to rank and distance the payoffs on separate dimensions rather than just on a dimension of overall satisfaction. He then weights the dimensions and combines the data through a form of geometric scaling.

d)?[9] The positive and negative signs would be the same as in figures 1A and 1B, but subjective monetary amounts would appear in the cells. This method may be all right for determining the relative distances between each payoff for a given judge, but it is not meaningful for making comparisons across judges because the amount of money a judge is willing to pay to avoid a dissatisfying alternative or to receive a satisfying one is partly dependent on how much a dollar is worth to him as well as how much dissatisfaction or satisfaction he feels from various outcomes. The idea of personally paying something no matter how small to avoid dissatisfaction or receive satisfaction may also seem too unrealistic for a judge to think about, although the approach may be made more manageable if the most satisfying alternative is automatically valued at $10 in order to provide a base line.

Another alternative to the type of questions posed in Figure 2 is simply to ask two questions instead of eight in order to arrive at a judge's threshold probability for cases in general or for a specific type of hypothetical case. The two questions would be as follows:

> 1. If you set a bond that *releases* a defendant prior to trial, he might subsequently fail to appear in court. This is undesirable result B. If you set a bond that *holds* a defendant prior to trial, he might have appeared in court if he would have been released. This is undesirable result A. Which of those two undesirable results would you consider more undesirable in the average case? A or B?
>
> 2. If we anchor the more undesirable of those two results at a minus 100 on a scale that goes from minus 100 to zero, then approximately where would you position the result that is not the more undesirable of the two? In other words, what relative number from —100 to 0 would you assign to B if you thought A was more undesirable, or to A if you thought B was more undesirable?

With those two items of information, we can now determine the

9. As an alternative based on these four questions, the judge might find it helpful to indicate first which possibility is the most satisfying, most dissatisfying, next to the most satisfying, and next to the most dissatisfying. He would then be asked, how many dollars of your own money would you be willing to pay to avoid the most dissatisfying possibility, to avoid the next most dissatisfying possibility, to receive the satisfaction of the most satisfying possibility, and to receive the satisfaction of the next to the most satisfying possibility? That two-stage approach (with four sub-parts to each stage) probably makes the evaluations easier although it does sometimes mean referring back to one's answers from the first stage in order to answer the second stage questions.

judge's threshold probability. Suppose, for example, the judge was William Blackstone who said that it is ten times as bad to convict an innocent person (a type A undesirable result) as it is to acquit a guilty person (a type B undesirable result).[10] Suppose further that Mr. Blackstone applied the same rule to holding and releasing in pre-trial release decisions. He would then in effect be saying that a type A result is worth 100 points on an undesirability scale and a type B result is worth 10 points on the same undesirabiliy scale.[11] Thus Mr. Blackstone would have a threshold probability of 10/(100 + 10) or .09, meaning he would release any defendant who has an appearance probability better than .09 and hold any defendant who has a probability less than .09. That .09 decision rule would in effect enable Mr. Blackstone to maximize his net satisfaction or expected values in his pre-trial release decision-making.[12] The advantage of

---

10. 4 W. BLACKSTONE, COMMENTARIES 358.

11. Since Blackstone and others tend to express the relative undesirability of a type 1 error to a type 2 error as a ratio, the questions in either the four-alternative form or the two-alternative form should probably also do so. This would involve, for example, wording the second question in the two-alternative form to read as follows:

> 2. How many more times as bad is result A over result B? (This assumes A was the more undesirable of the two results. Reverse the wording if B was mentioned as the more undesirable result.) In other words, is result A twice as bad as result B, five times as bad, ten times as bad, or is A some other multiple as bad as B?

The researcher can then give the more undesirable result a score of —100 and the less undesirable result a score of —100/M where M is the multiple indicating how many times more undesirable the worse alternative is compared to the less worse one. For example, if the respondent says it is five times as bad to have result A as result B, then he is in effect saying result A gets a score of —100 and B gets a score of —100/5, or —20. He is also in effect saying that his threshold probability is 20/(100+20), or .17.

12. The above approach assumes that a logically consistent decision maker would get an amount of satisfaction from avoiding a type 1 error equal to the amount of dissatisfaction from making a type 1 error. This means if the value of cell d in Figure 1 is found to be equal to —A, then the value of cell b is assumed to be equal to +A. Likewise, the above approach assumes that a logically consistent decision maker would get an amount of satisfaction from avoiding a type 2 error equal to the amount of dissatisfaction from making a type 2 error. This means that if the value of cell a in Figure 1 is found to be equal to —B, then the value of cell c is assumed to be equal to +B. Therefore, under the above approach, the expected value of releasing equals (+A)(PA) + (—B)(1—PA), and the expected value of holding equals (—A)(PA) + (+B)(1—PA). To find the threshold probability for a given judge involves setting those two expressions equal to each other and solving for PA. The result will be the equivalent of the algebraically simplified formula PA* = B/(A+B). It is also the equivalent of solving for PA in the formula previously given of (1—PA)(a) + (PA)b = (1—PA)(c) + (PA)(d) by substituting d, b, a, and c, for —A, +A, —B, and +B respectively.

An even simpler approach to determining one's threshold probability would be to use the formula PA* = 1/(X+1), where X = A/B. This approach merely involves determining the ratio between the amount of dissatisfaction from a type 1 error versus a type 2 error. It does this alternative is that it is simpler. Its disadvantage is that the respondent can thereby more easily see what he thinks are the socially acceptable answers and thus give those answers rather than his true attitudes.

One purpose for obtaining data like that asked for in Figure 2 would be to determine what kinds of background or attitudinal characteristics correlate with being a holding- or a releasing-oriented judge. That kind of information could be helpful in enabling persons involved in the judicial selection process to choose judges whom they consider as having a more appropriate or balanced orientation. Another purpose might be to provide the judges from a given court system with an analysis of how they stand on this pre-trial release dimension relative to their fellow judges so that they can decrease their releasing or holding orientation in order to come closer to the average judge in their system, or to come closer to a threshold of .50 or other threshold that might be considered desirable. Such a use would be analogous to informing the judges in a given court system how they compare in sentencing with the average judge in their system as is sometimes done among judges in order to produce more uniformity in their sentencing practices.

This uniformity-producing use of payoff matrices data logically raises the question as to the desirability of seeking uniformity among judges with regard to any PA* level or threshold probability of appearance other than .50. One might argue that if a defendant has a better than a .50 chance of appearing in court, then he should always be released; and if he has less than a .50 chance of appearing in court, then he should be held in jail pending a speedy trial. As previously implied, however, that reasoning may ignore the severity of the defendant's behavior and his likelihood of recommitting his crime before he is tried and sentenced, and it may ignore the relation within the court system between the cost of holding an average

---

not require determining the values of A and B but only the value of A/B. If that value is 10, as with Blackstone's standard for guilt, then PA* = 1/(10+1) = 1/11 = .09. That approach is probably the best approach for determining one's own threshold probability because of its simplicity, but it may not be a good approach to use in a questionnaire directed to judges or others because the respondents can too easily see what is involved, and thereby give answers that they think are socially desirable rather than their true answers. The approaches which seek values for A and B, or for a, b, c, and d are somewhat more complicated and time-consuming, but those defects may be more than offset by the increased subtlety and validity of those approaches. One can prove algebraically that 1/(X+1) is the equivalent of B/(A+B) and (a—c)/(a—b—c+d) given the definitions of these symbols.

defendant in jail and the cost of rearresting a released defendant. If the holding cost is substantially greater than the rearrest cost, then the system should be willing to release an average defendant even if his probability of appearing is substantially less than .50. Likewise, if the releasing cost is substantially greater than the average holding cost, then the system should be willing to hold an average defendant even if his probability of appearing might be greater than .50. State statutes specifying prerelease procedures normally allow discretion to deviate from a .50 figure, especially in view of the difficulty of determining what figure a judge is operating under.[13]

There is no way of determining the PA* threshold that a judge is using simply by observing his behavior. For example, if a judge releases 50 percent of two defendants, that judge may be operating at a .75 threshold level since he may have perceived one of the defendants as being above the .75 level and one as being below. On the other hand, that judge may be operating at a .25 level since he may have perceived one of the defendants as being above the .25 level and one as being below. In other words, by observing the percentage of defendants a judge releases, we cannot tell what his threshold probability-of-appearing criterion is unless we know what he perceived the probability of appearance figure to be for each of those defendants. If we had that information, we could observe above what PA figure he begins to release and below what figure (that is, the same figure unless there is a gray area) he begins to hold. With that information, we could assign each judge a behavioral PA* figure rather than just an attitudinal PA* figure which the questionnaire in Figure 2 generates although one's behavior generally follows one's attitudes.[14]

13. In order for a judge to have a .50 threshold PA* with +100 and —100 anchor points in the payoff cells, he would have to have the values of —100, +100, +100, and —100 in cells a, b, c, and d. No other combination of values could yield a .50 PA* with +100 and —100 anchor points except having all four cells be +100 or all four cells be —100, but it would be psychologically inconsistent for all four cells to be equally satisfying or equally unsatisfying. If having all judges operating with a .50 PA* were deemed desirable, then it would make sense to try to convince them that it is equally desirable to hold a defendant who would fail to appear in court as it would be to release a defendant who would appear in court. It would also then make sense to try to convince them that it is equally undesirable to release a defendant who had failed to appear in court as it would be to hold a defendant who would have appeared in court.

14. It might be difficult to devise a meaningful questionnaire or interviewing approach to determine in actual cases what a judge perceives the probability of appearing to be. This is so because it is quite possible that judges tend to reach an overall or holistic decision on

If we knew what a given judge perceived the probability of appearance figure to be for each defendant, then we might find that above a certain PA figure, he releases most, but not all defendants; and below that figure, he holds most, but not all defendants even though that figure is the one that involves the least inconsistencies. Those inconsistencies, however, may simply reflect the fact that the judge is not considering just one contingent event in making his release-hold decision. A second contingent event that he might be considering is the probability that the defendant may commit a crime while released although under most statutes, the probability of appearing in court is supposed to be the main or even exclusive criterion for releasing or holding defendants prior to trial. We could prepare a payoff matrix like those shown in Figure 1 in order to indicate how a given judge or type of judge feels in a given case or type of case about releasing or holding a defendant in light of the probability that he might commit a certain type of crime while released. From that four-celled matrix we could obtain for the judge a threshold PN* where PN stands for the probability of not committing a serious crime. For a defendant to be released under this multiple contingency perspective, he would have to have a probability of appearing greater than PA* and a probability of not committing a crime greater than PN*. If the defendant flunks either test, he does not get released.

3. MONETARY VALUES

One might ask how can the above multiple contingency approach take into consideration that more weight is supposed to be given to the probability of appearing in court than to the probability of committing a crime. This is difficult to do under the non-

whether to hold or release without making a determination (especially an explicit determination) of a defendant's probability of appearing. If judges were asked to write down what they perceived the PA to be in each case, they might have a tendency to say less than .50 where they had set a bond that resulted in holding the defendant and greater than .50 where they had set a bond that resulted in releasing the defendant. Perhaps some judges might be willing to record their perceived PA figures after observing and questioning the defendant, but before setting bond. This would be analogous to the cooperation the University of Chicago Jury Project received whereby judges agreed to indicate how they would decide a jury trial criminal case after the end of the evidence and the arguments, but before the jury reached its decision. H. KALVEN & H. ZEISEL, THE AMERICAN JURY 45-54 (1966). Another alternative research approach would be to present the judges with hypothetical situations like those discussed with regard to Figure 6 and ask them what they would estimate the probability of appearing to be in each situation given the information available.

monetary approach, partly because a —100 in the crime-committing payoff matrix is treated as being as undesirable as a —100 in the court-appearing matrix. Likewise, the non-monetary approach would treat a —100 in a payoff matrix dealing with murder crime-committing as being as undesirable as a —100 in a payoff matrix dealing with a substantially lesser crime. What we need is a common measurement unit for showing values across payoff matrices regardless of the contingent event with which we are dealing. Ideally, such a unit measures psychological utility, but that kind of unit is too difficult to express. As a substitute, we can at least tentatively try working with dollars.

Figure 3 shows a decision theory payoff matrix involving two contingent events and monetary values. The two contingent events are (1) appearing or failing to appear in court, and (2) committing or not committing a crime while released. The monetary values include two releasing costs and three holding costs. The releasing costs are (1) the cost of rearresting an average defendant, estimated at $200, and (2) the cost of a crime committed by an average defendant while released, estimated at $1,000. The holding cost consists of (1) the maintenance cost for keeping an average defendant in jail for an average pre-trial time period, figured at $4.43 per day for 2.28 months, (2) the lost gross national product for an average defendant, figured at $360 per month for 2.28 months, and (3) the bitterness cost that society might be willing to spend in order to avoid having a defendant sit in jail prior to trial for an average 2.28 months and then be found not guilty, estimated at $300 per month given that 12 percent of the defendants held in jail prior to trial are found not guilty.[15] Those figures come from a survey of police chiefs, prosecutors, defense attorneys, and bail administration officials in a sample of 23 cities, analyzed in more detail in another study.[16] The two releasing costs are symbolized RC_1 and RC_2 respectively in Figure 3A, and the three holding costs are collectively symbolized HC. Costs are shown as negative amounts, and benefits are shown as positive amounts. Releasing benefits (RB) are holding costs saved, and holding benefits (HB_1 and HB_2) are releasing costs saved.

15. Given the above figures, the average holding cost per defendant held is $1,206. Of that total, $303 is jail maintenance cost ($4.43 per day, times 30 days, times 2.28 months); $821 is lost gross national product ($360 per month, times 2.28 months); and $82 is bitterness cost ($300 per month, times 2.28 months, times 12 percent).

16. S. NAGEL, P. WICE, & M. NEEF, THE POLICY PROBLEM OF DOING TOO MUCH OR TOO LITTLE: PRE-TRIAL RELEASE AS A CASE IN POINT (1976).

FIGURE 3. DECISION-THEORY PAYOFF MATRICES INVOLVING TWO CONTINGENT EVENTS AND MONETARY VALUES

3A. RELEASING AND HOLDING COSTS

Alternative Possibilities with Point Probabilities

	(D) Would Fail to Appear & Commit Crime	(C) Would Appear & Commit Crime	(B) Would Fail to Appear & No Crime	(A) Would Appear & No Crime
	.08 X .15 = .01	.92 X .15 = .14	.08 X .85 = .07	.92 X .85 = .78
RELEASE	RB = +1,206 RC_1 = —200 RC_2 —1,000 $6	RB = +1,206 RC_2 = —1,000 $206	RB = +1,206 RC_1 = — 200 $1,006	RB = +1,206 $1,206
HOLD	HC = —1,206 HB_1 = +200 HB_2 = +1,000 —$6	HC = —1,206 HB_2 = +1,000 —$206	HC = —1,206 HB_1 = +200 —$1,006	HC = —1,206 —$1,206

3B. EXPECTED VALUES

Alternative Possibilities with Range Probabilities

	Would Fail to Appear & Commit Crime	Would Appear & Commit Crime	Would Fail to Appear & No Crime	Would Appear & No Crime	Total Expected Value
	.00 to .08	.07 to .15	.00 to .08	.77 to .85	
RELEASE	$.06 0 to .48	$29 14 to 31	$70 0 to 80	$941 (i.e., $1,206 × .78) 929 to 1025 (i.e., $1,206 × .77, etc.)	$1,040 943 to 1,136
HOLD	—$.06 —0 to —.48	—$29 —14 to —31	—$70 —0 to —80	—$941 —929 to —1025	—$1,040 —943 to —1,136

The same questionnaire data from which most of the cost figures were obtained also indicated that 92 percent of the released defendants appeared in court, and 85 percent of the released defendants were not known to have committed a crime while released.[17] Given that there are two contingent events, there are four possible occurrences which are labeled A, B, C, and D in Figure 3A. To determine the probability that a released defendant would both appear in court and not commit a crime, we could simply multiply the .92 by the .85 if we are willing to assume that those two subpossibilities are independent of each other. We know, however, that they are not likely to be completely independent of each other because defendants with certain characteristics are likely to both appear and not commit crimes, whereas defendants with opposite characteristics are more likely to both fail to appear and to commit crimes. Even though we do not know how closely related our two contingent events are, we can from the data we have determine a maximum and a minimum probability somewhere between .00 and 1.00 for the A, B, C, and D alternative possibilities. This is done in Figure 4. For example, if possibilities A and C must constitute 92 percent of our released defendants (since those two possibilities constitute all the defendants who appear and only the defendants who

17. Just because 92 percent of the released defendants appear in court does not mean that 92 percent of all the defendants would appear in court since the 27 out of 100 defendants who are held do not get an opportunity to fail to appear. Assuming their failure rate is about double the 8 percent rate for those who are released, this means that if all the defendants were released, then the 92 percent appearance rate would drop to 90 percent. This figure is arrived at by weighting the .92 appearance rate by the fact that it covers .73 of the defendants, and by weighting the .84 appearance rate (i.e., 100 percent minus 16 percent) by the fact that it covers .27 of the defendants. In other words, the new .90 appearance rate equals (.73)(.92) + (.27)(.84). If we assume the failure rate for the detainees is triple the rate for those released, then we would calculate (.73)(.92) + (.27)(.76) which yields an appearance rate of .88. If we go so far as to assume none of the detainees would appear if released, then the overall appearance rate would still be .67 or two-thirds since (.73)(.92) + (.27)(0) = .67.

The reason the appearance rate is not drastically changed by figuring in those who are not released is because (1) almost three-fourths of the defendants are released, (2) only a low 8 percent are known to fail to appear, and (3) because many defendants who are not released may be good risks but they lack the funds to pay the bond or they are possibly misperceived as being bad risks. Thus with this data, each doubling of the failure to appear rate of those not released only results in a reduction of 2 percent in the general appearance rate since .73 times .08 is .02. Even if a substantially lowered appearance rate figure were used, the general conclusions of this section would not be changed that the expected value of releasing the average defendant is substantially greater than the expected value of holding the average defendant. Another useful aspect of this kind of analysis is that one can easily change the probabilities (or the costs) and see how the results would change with regard to which of the alternative decisions is the best decision in benefit-cost terms, as in note 23 *infra*. The same above considerations apply to the .85 tentative probability which is used to indicate the percent of the released defendants who are not known to have committed a crime while released.

appear), then that equality alone indicates that either possibility A or possibility C (but not both) can have a maximum probability of .92. The second equality tells us, however, that possibility A can be no higher than .85 and not .92. Applying that same reasoning to A, B, C, and D in Figure 4A, we can then take those four maximum probabilities and the four equalities to Figure 4B to determine what the minimum probabilities have to be. Thus, A could not be lower than .77. Otherwise, A + C could not add up to .92, since .15 is as high as C can be.[18]

Now that we have that cost and proability data, what do we do with it? That is where Figure 3B comes in. The logical thing to do is to determine for each of the four alternative possibilities what the expected values are for releasing or holding this hypothetical average defendant. The expected value for any one of the eight cells in our payoff matrix is the total value or cost of that cell times the probability that the combination of contingent events will occur

18. If we had 100 cases of defendants who were released, the following four-cell table would be consistent with our data that shows 92 percent of the released defendants appear in court and 85 percent of the released defendants do not commit crimes.

APPEARANCE

		Fail	Appear	
CRIME-COMMITTING	No Crime	B 5	A 80	85
	Crime	D 3	C 12	15
		8	92	100 cases

With that hypothetical data, the probabilities of occurrences A, B, C, and D are .80, .05, .12, and .03 respectively, all of which are consistent with the ranges given in Figure 3B. These are referred to as empirical combined probabilities, rather than empirical single probabilities like the .92 probability of appearing, or *a priori* combined probabilities like the .92 times .85 probability of appearing and not committing a crime. With that hypothetical data, one can also say that if a defendant appears, the probability that he did not commit a crime is 80/92 or .87. Likewise, if a defendant does not commit a crime, the probability that he will appear is 80/85 or .94. Knowing any one of the four categories on the columns or rows, we can give a probability for any of the other categories. These are referred to as conditional or Bayesian probabilities. If we know with 1.00 accuracy that a defendant failed to appear, then we know of course that there is a zero probability that he appeared since those are complimentary probabilities.

FIGURE 4. DETERMINING THE MAXIMUM AND MINIMUM PROBABILITY FOR EACH CONTINGENT POSSIBILITY WITH TWO CONTINGENT EVENTS

4A. DETERMINING THE MAXIMUM PROBABILITY FOR EACH POSSIBILITY

Given this Equality	Then Max D is	Then Max C is	Then Max B is	Then Max A is
A + C = .92	X	.92	X	.92
A + B = .85	X	X	.85	<u>.85</u>
C + D = .15	.15	<u>.15</u>	X	X
B + D = .08	<u>.08</u>	X	<u>.08</u>	X

(The lowest maximum for each possibility is underlined.)

4B. DETERMINING THE MINIMUM PROBABILITY FOR EACH POSSIBILITY

Given this Equality and the Above Maximums	Then Min D is	Then Min C is	Then Min B is	Then Min A is
A + C = .92	X	.07 (*i.e.*, .92-.85)	X	.77 (*i.e.*, .92-.15)
A + B = .85	X	X	.00 (*i.e.*, .85-.85)	.77 (*i.e.*, .85-.08)
C + D = .15	.00 (*i.e.*, .15-.15)	.07 (*i.e.*, .15-.08)	X	X
B + D = .08	.00 (*i.e.*, .08-.08)	X	.00 (*i.e.*, .08-.08)	X

(An X means the equality does not apply.)

which that cell represents. Thus, the expected value for releasing an average defendant who would fail to appear and not commit a crime (possibility B) is \$70. That figure represents the RB + RC_1 sum (or \$1,006) multiplied by the .07 probability. If one rejects the above independent probability approach for handling two contingent events, one can say the expected value is from \$0 to \$80. That range represents the RB + RC_1 sum multiplied by the .00 minimum probability, and the RB + RC_1 sum multiplied by the .08 maximum probability for alternative possibility B.[19]

After we have obtained expected values for each payoff cell, the next logical thing to do is to sum the expected values across each row in order to obtain the total expected value for releasing the average defendant in order to compare that value with the total expected value for holding the average defendant. Doing so with the data given shows that our judicial system would maximize its benefit-cost picture by releasing the average defendant. This is what our judicial system does do, at least as indicated by the same questionnaire data from which the costs and probabilities were taken. That data shows the average defendant has a .73 chance of being released since the average city in the sample reported a 73 percent release rate. Many defendants, however, are not average defendants which possibly explains why 27 percent of the defendants are held in jail pending trial. Perhaps the judges in those specific cases implicitly perceive the probability of appearing, the probability of crime committing, the releasing costs, and the holding costs to be such that the expected value of holding those defendants is greater than the expected value of releasing them rather than the more common reverse order.[20] Perhaps some of those cases also involve other contingent events (like the probability that a defendant al-

19. Essentially the same information presented by the eight-cell tables of Figure 3 could be presented by a decision tree that ends in eight branches. It begins with the decision fork of releasing or holding, with each of those two forks going into an appearing versus a failing-to-appear fork, and each of those forks going into a no-crime versus a crime fork. A decision tree approach, however, consumes more space to present, conveys less detail, contains more repetitive labels, has extra arithmetic steps by not working with combined probabilities, and is probably not so easy to read. A decision tree approach is also arbitrary in whether it presents the appearance contingency before or after the crime-committing contingency since they are not sequential events. The results, however, should be identical between the tabular approach and the decision tree approach since the two approaches are basically just different methods of visual presentation, not differences in substance.

20. Prediction techniques which consider the characteristics of the defendant and his crime can be useful in predicting his probabilities and costs. The most widely known prediction scheme for predicting the probability of appearing was developed by the Vera Institute in New York City. See Ares, *supra* note 2. For predicting the probability of crime-committing while released, see J. Locke, *supra* note 3. On the use of subjective probabilities rather than probabilities based on statistical data, see Huber, *supra* note 7, and KOTLER, *supra* note 7. On the use of statistical techniques for arriving at probabilities, see S. NAGEL, THE LEGAL PROCESS FROM A BEHAVIORAL PERSPECTIVE 144-72 (1969); and D. FINNEY, PROBIT ANALYSIS (1971). For a recent example of the application of probit analysis (which involves using statistical techniques to arrive at probabilities), Warren Hausman and Richard Thaler of the University of Rochester School of Management have been experimenting with the application of probit analysis for obtaining pre-trial release prediction probabilities.

though "guilty" will be acquitted) or other costs (like the opportunity cost of not taking advantage of the opportunity of holding a defendant in order to teach him and others a lesson not to get arrested).

The decision rule generated by an analysis of Figure 3 is: Release a defendant when the expected value of releasing him given his specific probabilities and costs is greater than the expected value of holding him. An alternative way of conceptualizing the release-hold decision would be to say: Release the defendant if the expected holding costs are greater than the expected releasing costs. The expected holding costs for our hypothetical average defendant are $1,206 or HC, assuming he is held. The expected releasing costs are $(PF)(RC_1) + (PC)(RC_2)$ where PF is the probability of failing to appear (i.e., 1.00 — PA) and PC is the probability of committing a crime while released (i.e., 1.00 — PN). The expected releasing costs are thus (.08)($200) + (.15)($1,000), or $16 + $150, or $166. That alternative conceptualizing could also be stated as: Release the defendant if the expected releasing benefits are greater than the expected holding benefits. Given the data, the expected releasing benefits for the average defendant would be $1,206 saved, and the expected holding benefits for the average defendant would be $166 saved.[21] We can also combine those two alternative conceptions by saying release the defendant if releasing benefits minus releasing costs (i.e., $1,206 minus $166) is greater than holding benefits minus holding costs (i.e., $166 minus $1,206). In other words, the $1,040 expected value of releasing for the average defendant is greater than the —$1,040 expected value of holding.[22] Either the expected value

21. It is *not* meaningful to say that someone released who appears in court without committing a crime has provided the system with $200 in benefits by not having to be rearrested or with $1,000 in benefits by not having committed a crime. All one can say is that such a person has not caused the system to incur $200 in rearrest costs ($RC_1$) and has not caused the system to incur $1,000 in crime-committing costs (RC_2).

22. A related conclusion is reached in Friedman, *The Evolution of a Bail Reform*, 7 POLICY SCIENCES 281 (1976). In his appendix, Friedman shows that the benefits of releasing on recognizance are greater than the costs of releasing on recognizance, at least in New York City. He does *not* do so by working with monetary values for all the benefits and costs involved, but only for those that are relatively easy to assign a monetary value to. The only holding cost he deals with is the jail maintenance cost which he figures at $3 a day for an average of 30 days per defendant held. The only ROR releasing cost he deals with is the cost of interviewing, verifying, and following up on released defendants which he figures at $45 per defendant released. Thus, the problem for him is whether $90 + V is greater than $45 + .016C. The V represents the benefits from releasing other than the $90 saved in jail mainte-

approach via the payoff matrix or the expected net benefit approach (i.e., expected benefits minus expected costs) should yield the same result which in this case is a decision in favor of releasing the average defendant prior to trial.[23]

---

nance. The C represents the costs of releasing other than the \$45 spent as par of the ROR processing. The .016 indicates that those costs are only incurred in the 1.6 percent of the cases where the ROR defendant fails to appear.

He establishes that \$90 + V must be greater than \$45 + .016C by the following steps:

(1) Since the average defendant is released rather than held, this means 90 + V must be valued more than .04C, where .04 indicates that 4 percent of the released defendants failed to appear (but not released ROR).

(2) Then C < 2,250 + 25V, by interchanging both sides of the inequality and dividing both sides by .04.

(3) Then 45 + .016C < 45 + .016 (2,250 +25V), by substituting 2,250 + 25V as a value greater than C.

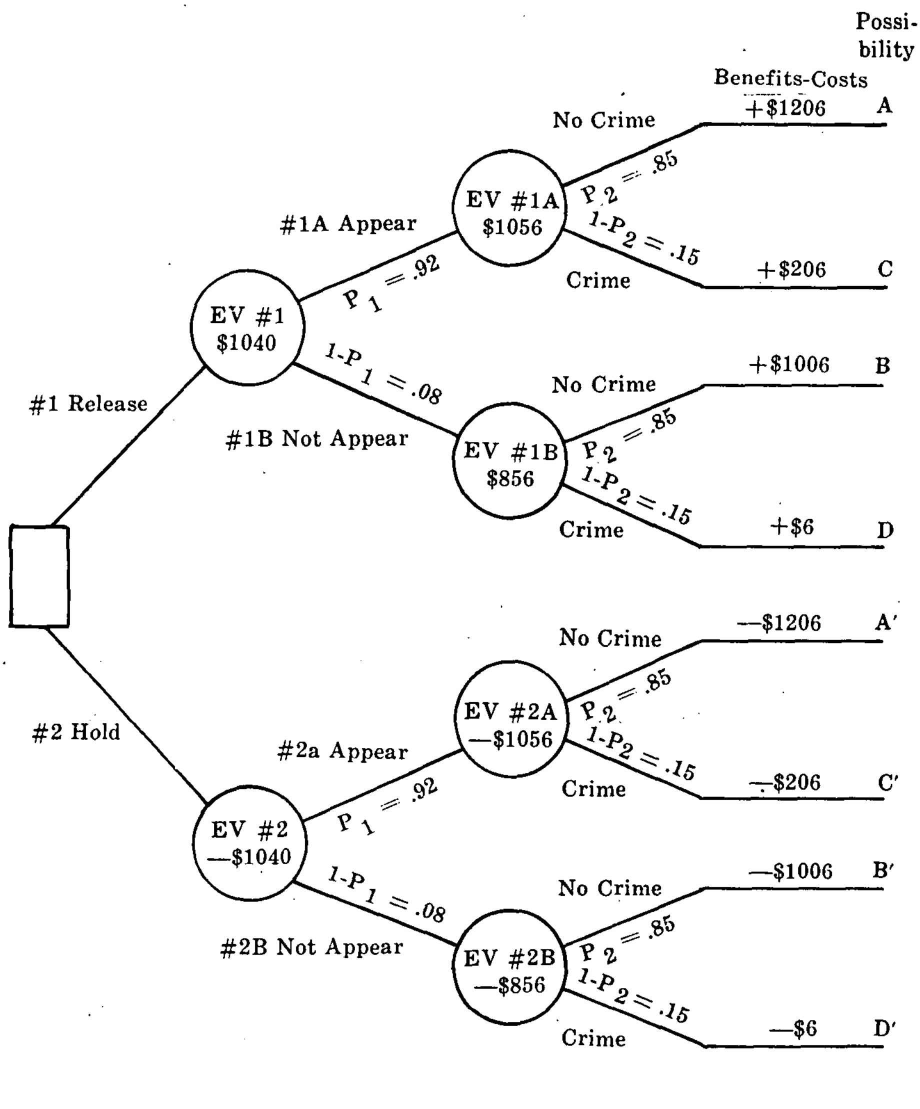
(4) Then 45 + .016C < 81 + .4V, by simplifying the right side of the inequality through the removal of the parentheses.

(5) Therefore, the costs of ROR (i.e., 45 + .016C) are less than the benefits of ROR (i.e., 90 + 1V) since 81 + .4V is less than 90 +1V. If d < e and e < f, then d < f and f > d.

In other words, what Friedman is basically saying is that society must consider releasing to be more valuable than holding since society does more releasing than holding, and therefore it must consider ROR to be more valuable than holding since ROR is a form of releasing that has a better appearance rate than releasing in general. The big defect in his analysis is that he does not show that releasing or ROR is actually more profitable than holding, but rather he only shows that society must perceive releasing or ROR as being more profitable if one operates on the assumption that society or an individual decides in favor of the alternative activity that is perceived to be the most profitable activity.

23. The same analysis based on Figures 3A and 3B and the above calculations could be applied to determining the sensitivity of the outcome to changes in the inputs. In other words, how much would the inputs have to be changed in order to reverse the decision from favoring the release of the defendant to favoring his being held? More specifically, how much would the .92 appearance rate have to drop before the expected value of releasing would fall below the expected value of holding? If we were only concerned with one holding cost and one releasing cost and only with the probability of appearing rather than the probability of crime committing, then the formula for the expected value of releasing would be (1—PA)(RB—RC) + (PA)(RB). In other words, the value of cell a in Figure 1 is the releasing benefits minus the releasing costs or RB—RC, and the value of cell b is RB as is shown below. Likewise, the formula for the expected value of holding would be (1—PA)(HB—HC) + (PA)(HC), which means the value of cell c in Figure 1 is HB—HC and the value of cell d is HC. Now all we have to do to answer the above question (as to how low does PA have to drop to make the expected value of holding greater than the expected value of releasing) is (1) substitute monetary values for RB, RC, HB, and HC, (2) set those two formulas equal to each other, and (3) solve for the value of PA. Any probability lower than that value would make it more worthwhile to hold the defendant. Actually with the releasing and holding cost data for the average defendant from Figure 3, it would always be more profitable to release such a defendant no matter how low the value of PA is. This is so because all the cell values on the releasing row are higher than the corresponding cell values on the holding row which they would not be if the holding costs were lower or the releasing costs were higher. By using the

FIGURE 5. DECISION TREE INVOLVING TWO CONTINGENT EVENTS AND MONETARY VALUES



$EV\#1A = (P_2)(A) + (1\text{-}P_2)(C)$ $EV\#1B = (P_2)(B) + (1\text{-}P_2)(D)$

$EV\#2A = (P_2)(A') + (1\text{-}P_2)(C')$ $EV\#2B = (P_2)(B') + (1\text{-}P_2)(D')$

$EV\#1 = (P_1)(EV\#1A) + (1\text{-}P_1)(EV\#1B)$

$EV\#2 = (P_1)(EV\#2A) + (1\text{-}P_1)(EV\#2B)$

Another insight-generating way of analyzing the data which was placed in Figure 3 is to use a decision tree perspective like that shown in Figure 5. The left trunk end of the decision tree shows that we are trying to determine the value of a decision to release versus a decision to hold. Releasing can result in appearance or non-appearance and in no crime or a crime. Likewise, the held defendant could have resulted in appearance or non-appearance and in no crime or a crime if the defendant would have been released. At the right end of each branch are given the A, B, C, or D possibilities listed in Figure 3, along with the benefits minus costs from Figure 3. A prime sign is used to distinguish those possibilities for held defendants versus released defendants. The expected value of any branch that is not an end branch is equal to the sum of the values

---

data in Figure 3 and solving for PA* by the above approach or the formulas given in footnotes 6 or 12 one obtains a PA* that is negative but rounds off to zero which is the nearest possible PA.

	Fail to Appear (1-PA)	Appear (PA)	Expected Value
Release	RB - RC	RB	(1-PA)(RB-RC) + (PA)(RB)
Hold	HB - HC	HC	(1-PA)(HB-HC) + (PA)(HC)

One could similarly answer the question, how much would the holding costs have to decrease (or the releasing costs have to increase) in order to make the expected value of holding greater than the expected value of releasing? To answer that question about HC, one would substitute numerical values for PA and RC. This would also give us the numerical value of HB since holding benefits are simply the positive sign of the minus releasing costs. We could then use —X to label HC and X to label RB. The next step would be to set those two formulas equal to each other as we previously did, but this time solve for X rather than PA. If the holding costs decrease $1 below the value of X, then it is more worthwhile to hold the defendant than to release him assuming everything else is held constant. With the partial data from Figure 3, the value of X is only $16 since the expected value of releasing is (.08)(X—200) + (.92)(X), and the expected value of holding is (.08)(—X+200) + (.92)—(X). This means the holding cost would have to be drastically reduced (or the releasing cost drastically raised) in general or in a specific case to make holding more worthwhile than releasing if one only considers the probability and cost given of rearresting for non-appearance. One could easily extend the above sensitivity analysis (i.e., the sensitivity of the result to changes in the inputs) by expanding the formulas to include more kinds of holding and releasing costs and more kinds of contingent probabilities analogous to those in Figure 3.

of the subsequent branches discounted or multiplied by the probability of their occurring, as indicated by the equations given at the bottom of the figure. Those calculations are referred to as folding back because they involve working backwards from the values on the horizontal end branches and the probabilities on the diagonal branches to the expected values in the previous circles.

The decision tree shows that the expected value of releasing an average defendant is substantially greater than the expected value of holding one as was previously indicated. That result is not affected by whether the decision tree branches first on appearance and then on crime-committing, or first on crime-committing and then on appearance. That result, however, assumes that the probabilities of appearing and crime committing are not affected by whether we release or hold someone. In other words, if an average defendant has a .92 probability of appearing, we assume he would have had that same probability of appearing if released regardless whether he is actually released or held. Likewise, the model assumes the benefits and costs are not affected by whether we hold someone. In other words, if an average defendant who commits a crime while released incurs $1,000 in social costs, he would incur the same $1,000 if he were to commit a crime while released regardless whether he is released or held. A decision tree perspective can often be quite helpful in analyzing decisions that involve more than one contingent event, more than one decision-branching point, or more than dichotomous branches.

B. *The Bond-Setting Decision*

So far, we have been discussing the decision problem of just whether to release or hold a defendant prior to trial. That, however, is not the way the pretrial release problem is usually stated in the courtroom context. In that context, the arraignment judge is usually faced with the decision of what bond to set for the defendant rather than the dichotomous decision of whether or not to release the defendant. Nevertheless, to a considerable extent, the bond-setting decision can be reduced to a release-hold decision if one equates "release" with a low bond (that is, a bond the defendant can and will meet) and "hold" with a high bond (that is, a bond the defendant cannot or will not meet). If an arraignment judge wants to release a defendant, but the defendant is unexpectedly unable to meet the bond the judge has initially set, then the judge can lower it to arrive at a figure the defendant can meet. On the other hand, if an arraignment judge wants to hold a defendant and he sets a very high bond which the defendant unexpectedly can meet, then the judge cannot so easily raise the bond without some evidence of changed circumstances other than the circumstance that the defendant had more money than the judge thought he had.

One interesting aspect about the bond-setting decision as contrasted to the releasing decision is that the decision itself can influence the probability of appearing in court, which is supposed to be the main criterion in arriving at the decision. Thus, a judge can increase the probability of a defendant appearing in court by setting a high bond, provided that the defendant can meet the bond. With a high bond, the defendant has more of an incentive to appear in order to retrieve his bond than he would with a low bond. A rational way to combine the bond-setting decision with the decision to release or hold might involve a five-step process. First, the judge involved can determine his own threshold probability of appearance (PA*) through the payoff matrix approach we discussed in Section II.A or through whatever method the judge prefers. Second, the judge can determine whether the defendant's probability of appearing in court is greater than that threshold probability regardless of the bond set. Third, if the defendant's PA is greater than PA*, then the defendant can be released on his own recognizance or on a nominal bond. Fourth, if the defendant's PA is less than PA*, then the bond should be set just high enough to bring the defendant's PA above the threshold probability. Doing so involves the judge taking into consideration the defendant's ability to pay since a low bond will be more of an incentive for a poor person to appear in court than a rich person.[24] Fifth, if no bond can bring the defendant's probability of appearing above the threshold, then the defendant may have to be held in jail pending a speedy trial. Likewise, if the bond that

24. Somewhat contrary to this rational scheme, however, is the fact that bond tends to be set lower for non-indigent defendants than for indigent ones if we count release on recognizance as a zero bond. The median or middlemost bond for a nationwide sample of 246 indigent defendants charged with grand larceny was $2,328, whereas the median bond for a similar sample of 354 non-indigents was only $1,850. Further details are given in Nagel, *Effects of Alternative Types of Counsel on Criminal Procedure Treatment*, 48 IND. L.J. 404 (1973), especially notes 15 and 28 and the accompanying text. That seeming discrepancy can, though, be partly explained by the possible fact that non-indigent defendants have a greater probability of appearing in court than indigent defendants do, especially if there is no trial-day notification system.

can bring PA above PA* is too high for the defendant to meet, then he may also have to be held while awaiting trial. It is unfortunate when a defendant has to be held in jail pending trial given the fact that the holding costs are usually higher than the releasing costs. That fact, however, is implicitly taken into consideration in the lowness of the threshold probability which the judge uses as his criterion in determining which defendants to hold.

The bond-setting decision, like the releasing decision, can simultaneously consider the two contingent events of appearing in court and not committing a crime while released. In that decision situation, our five-step decision process would be adjusted as follows. First, the judge involved determines his PA* threshold with the court-appearance contingency, and then his probability-of-not-committing-a-crime threshold (PN*) with the crime-committing contingency. Second, the judge determines whether the defendant's probability of appearance is greater than PA*, and whether the defendant's probability of not committing a crime (PN) is greater than PN*. Third, if the defendant passes both tests, then he can be released on his own recognizance or on a nominal bond. Fourth, the judge can try to set a bond high enough to bring PA over PA*. Doing so, however, may sometimes decrease PN since a high bond has been known to motivate a defendant to commit a crime in order to pay off the high bond loan or high bond premium. Fifth, the defendant may have to be detained in jail until his trial if step four does not bring both his probabilities above the threshold cutoffs.

As an alternative to the above five-step process, a judge could conceivably go through a kind of crude or implicit expected value calculation with monetary values like those shown in Figure 3. The only difference would be that instead of the alternative decisions being release or hold, they would be (1) release on the highest bond that the defendant can meet, or (2) hold. The probabilities, with regard to appearing in court at the top of columns A and C would thus be higher than if the choice were merely release or hold, and the probabilities with regard to failing to appear at the top of columns B and D would be lower. All the other calculations would be the same, and the defendant would be released if the expected value of releasing (EV_R) were greater than the expected value of holding (EV_H).

Perhaps we should emphasize that we are not saying that judges do or should prepare payoff matrices for each arraignment case. What we are saying is that the payoff matrix approach can provide an understanding of what may be implicitly happening in an inexact way in a judge's mind. What we are also saying is that the payoff matrix or decision theory approach can provide a means for understanding the effects of various cost changes and probability changes on releasing and bond-setting decisions. That approach, when combined with a questionnaire or other data-gathering techniques, can also provide information relevant to encouraging more uniformity or more compliance with appropriate legal standards in releasing and bond-setting decisions.

A good set of questions to add to our Figure 1 questionnaire might especially include some short hypothetical bond-setting problems. For example, the judges might be presented with 15 hypothetical defendants and asked to place each one in about a dozen bond categories. Four such hypothetical case questions are shown in Figure 6. That particular set is designed to determine the extent to which a responding judge is influenced by the severity of the crime committed irrespective of its relation to the probability of the defendant's appearing in court. Judges will generally admit they set bond higher for more severe crimes but not in order to punish the defendant with pretrial detention, but rather in recognition of the relation between crime severity and failure to appear in court.[25] If a judge gives a high bond in cases 4 and 7 and a low bond in cases 1 and 15, then he is more influenced by the probability of appearance than by the charge. If a judge gives a high bond in cases 1 and 7 and a low bond in cases 4 and 15, then he is more influenced by the charge than by the probability of appearance. If a judge is about equally high or low in all four cases, then he is not influenced by either the severity of the charge or PA. It might be interesting to know what explains those different propensities among different judges. It might also be interesting to point out their propensities to them in an unpublicized way possibly as part of a judicial workshop to see if doing so might change their subsequent bond-setting behavior in court. This questionnaire approach obviously works better than directly asking a judge what his propensities are, and works better than trying to determine his propensities from a mass of cases in which the possibly influential variables cannot be separated out or controlled for. Other hypothetical defendants can be included to

25. P. WICE, *supra* note 4, at 25-34.

get at the role of the probability of crime committing, the effect of economic class, and to decrease the visibility to the respondent of the comparisons that are likely to be made.

FIGURE 6. SOME QUESTIONS FOR OBTAINING A JUDGE'S PROPENSITIES REGARDING BOND-SETTING DECISIONS

General Statement: Listed below are 15 defendants appearing before you in an arraignment proceeding. For each defendant, indicate in what range you would generally set bond for an average defendant having those characteristics. The ranges we are using are as follows: (1) release on recognizance; (2) \$0 to \$99; (3) \$100 to \$499; (4) \$500 to \$999; (5) \$1,000 to \$1,999; (6) \$2,000 to \$4,999; (7) \$5,000 to \$9,999; (8) \$10,000 to \$19,999; (9) \$20,000 to \$49,999; (10) \$50,000 to \$99,999; (11) \$100,000 or over; (12) no bond allowed.

Item #	*Charge*	*Probability of Appearing in Court*	*Other Characteristics*	*Bond Category*
1.	Armed robbery	.90	None available	______ (Indicate category 1 to 12)
.				
.				
.				
4.	Shoplifting	.20	Adult male	______
.				
.				
.				
7.	Armed robbery	.15	None available	______
.				
.				
.				
15.	Shoplifting	.85	Adult male	______

C. *Increasing the Expected Value of Releasing in Individual Cases*

1. THE GENERAL PERSPECTIVE

In an arraignment proceeding, a judge must decide whether to release the defendant on his own recognizance or by way of a low bond, or whether to hold the defendant by refusing to set bond or by way of a high bond. Judges are probably more reluctant to make an error of releasing a defendant who would fail to appear (a type 2 error) than to make an error of holding a defendant who would appear if released (a type 1 error). Releasing errors are more avoided because they are more visible than holding errors since it is embarrassing to a judge if a defendant he released fails to appear, but no one knows for sure if a defendant he held would have appeared if he would have been released. As a result there may be much more pretrial holding than is necessary.

What is needed is to make the holding errors and the holding costs more visible in order to decrease the unnecessary and wrongful holding. What may also be needed is to decrease the releasing costs so that judges will be more willing to release defendants. In addition, there is a need for raising and clarifying the probability that defendants will appear in court since much of the holding may be based on (1) an unduly low actual or perceived probability of appearing, and also on (2) an unduly vague probability of appearing plus an implicit strategy saying to hold rather than release when the situation is unclear. The tendency to hold when in doubt and the greater sensitivity to releasing errors rather than holding errors run contrary to the rule of law in pretrial release decisions which says the benefit of the doubt concerning appearance should go to the defendant as part of the general presumption of innocence and the presumption of appearing in court.[26] Thus, we are talking about promoting the rule of law in pretrial release decisions when we talk about the above-mentioned need for raising and clarifying the probability of appearance, making more visible the type 1 errors and costs of holding defendants who would appear, and decreasing the costs of type 2 errors of releasing defendants who fail to appear.

In the top row of the four-cell table in Figure 7, we show that the expected value of releasing (rather than holding) a defendant logically equals (1) the benefits of releasing minus (2) the costs of releasing. The benefits are the positive value to the judge of having the released defendant appear in court (symbolized +A), and the costs are the negative value to the judge of having the released

26. The ABA MINIMUM STANDARDS FOR CRIMINAL JUSTICE RELATIVE TO PRE-TRIAL RELEASE provides, at section 1.1, "The law favors the release of defendants pending determination of guilt or innocence," and at section 5.1 it provides, "It should be presumed that the defendant is entitled to be released in order to appear on his own recognizance." The Illinois Criminal Code, 38 ILL. ANN. STAT. § 110-2 (1963) provides:

> When from all the circumstances the court is of the opinion that the accused will appear as required either before or after conviction, the accused may be released on his own recognizance . . . . This Section shall be liberally construed to effectuate the purpose of relying upon criminal sanctions instead of financial loss to assure the appearance of the accused. FED. R. CRIM. P. 46.

18 U.S.C. §§ 3146(a) and (b) provide that:

> A person arrested for an offense not punishable by death shall be admitted to bail . . . . the amount thereof shall be such as in the judgment of the commissioner or court or judge or justice will insure the presence of the defendant, having regard to the nature and circumstances of the offense charged, the weight of the evidence against him, the financial ability of the defendant to give bail, and the character of the defendant.

defendant fail to appear in court (symbolized —B). We cannot, however, merely determine A minus B, because doing so would not take into consideration that those benefits and costs are contingent on the probability of the defendant appearing in court. Instead, the benefits have to be discounted by the probability of their occurring (that is, the value of A multiplied by P), and the costs have to be discounted by the probability of their occurring (that is, the value of B multiplied by 1 minus P).

Likewise, the expected value of holding a defendant in the bottom row of the table equals (1) the discounted benefits of holding minus (2) the discounted costs of holding. The benefits are the positive value to the judge of having held the defendant when he would have failed to appear in court if released (symbolized +B), and the costs are the negative value to the judge of having held the defendant when he would have appeared in court if released (symbolized —A). The expected value or discounted net benefits is thus equal to $(-A)(P) + (+B)(1-P)$. Holding a defendant who would have appeared in court is a type 1 error (or an alpha error, which explains the use of the letter A), and releasing a defendant who would fail to appear is a type 2 error (or beta error, which explains the use of the letter B). A type 1 error rejects the hypothesis (or presumption that the defendant will appear) when it is true, and a type 2 error accepts the hypothesis (that the defendant will appear) when it is false.

As one can see by some simple algebraic manipulation, the expected value of releasing varies (1) directly or positively with the values of P and A, and (2) inversely or negatively with the value of B. Likewise, the expected value of holding varies (1) inversely with the values of P and A, and (2) directly with the value of B. Thus, if we want to increase the expected value of releasing relative to the expected value of holding, we should logically seek to increase P, increase A, and decrease B. In other words, we should seek to influence the awareness, perceptions, and facts which relate to (1) the probability of a defendant appearing in court, (2) the errors and costs of his being held when he would have appeared, and (3) the errors and costs of his failing to appear when released.[27]

27. An alternative way to categorize what needs to be done to increase the expected value of releasing relative to the expected value of holding might involve thinking in terms of four separate cells, rather than two key cells where those two cells determine the values of the

FIGURE 7. INCREASING THE PERCENTAGES OF DEFENDANTS RELEASED IN THE PRETRIAL RELEASE DECISION

PROBABILITY OF APPEARANCE

		Would Fail to Appear (1-P)	Would Appear (P)	EXPECTED VALUES
ALTERNATIVE DECISIONS AVAILABLE	Release via ROR or Low Bond	Type 2 Error $-B$	$+A$	$EV_R = (-B)(1\text{-}P) + (+A)(P)$
	Hold via No or High Bond	$+B$	Type 1 Error $-A$	$EV_H = (+B)(1\text{-}P) + (-A)(P)$

There are three general approaches to widening the positive difference between EV_R and EV_H :

I. Raise and clarify the probability of appearance (i.e., Increase P).
   1. Raise P through better screening and notification.
   2. Clarify P through statistical studies of what percentage of various types of released defendants appear in court.
   3. More vigorously prosecute those who fail to appear.

II. Make more visible the type 1 errors and costs of holding defendants who would appear (*i.e.*, Increase A).
   1. Publicize for each judge the percent of defendants he holds and the appearance percent he attains. (Judges vary widely on percent held, but appearance percentages tend to be about 90 percent.)
   2. Make more visible how much it costs to hold defendants in jail.
      a. Jail maintenance
      b. Lost income
      c. Bitterness from case dismissed after lengthy wait
      d. Families on welfare
      e. Increased conviction probability
      f. Jail riots from overcrowding

III. Decrease the costs of type 2 errors of releasing defendants who fail to appear (*i.e.*, Decrease B).
   1. Make rearrest more easy through pretrial supervision.
   2. Decrease the time from arrest to trial.
      a. More personnel, more diversion, and shorter trial stage.
      b. Better sequencing of cases.
      c. Shorter path from arrest to trial.
   3. Decrease pretrial crime committing.
      a. Increase probability of being arrested, convicted, and jailed.
      b. Decrease benefits of successful crime committing.
      c. Increase costs of unsuccessful crime committing.

2. RAISE AND CLARIFY THE PROBABILITY OF APPEARANCE

An important way to increase the probability that released defendants will appear is through better screening and notification. Better screening might involve a combination of the subjective analysis of a defendant's probability of appearing plus a more objective analysis like that used in the point system of the Vera Institute. Notifying the defendant by postcard or telephone the day before his court appearance may often prevent non-appearances especially by low income defendants who are not accustomed to middle class procedures for appointment keeping.

To improve the effectiveness of pretrial release screening, more research is needed to determine what set of variables are the best predictors and what statistical procedure is the best way to weight the variables. A study in Rochester interestingly revealed that one can get better predictability from checking whether or not the defendant was telling the truth with regard to how long he has held his present job. This yes-no variable predicts better than the variable of whether the defendant has held his present job one month, six months, a year, or longer.[28]

These statistical techniques, by aiding in the screening process, improve the probability that released defendants will appear. They also clarify that the appearance rate is high for defendants in certain categories and for defendants in general. That clarification can increase the perceived probability of appearance even if it does not affect the actual probability of appearance. This is important since judges are acting on the basis of their perceptions of P although that indirectly is influenced by what P actually is.

---

other two cells although they are opposite in sign. More specifically, if releasing is considered the more desired behavior and holding the less desired behavior, then what needs to be done is to (1) increase the (perceived) benefits of releasing, (2) decrease the costs of releasing, (3) decrease the benefits of holding, and (4) increase the costs of holding. Item 2 in this list of four items corresponds to decreasing B, and item 4 corresponds to increasing A. Item 1 mainly involves increasing the satisfaction that comes from saving the holding costs, which is the equivalent of making more visible the value of A. Item 3 mainly involves decreasing the satisfaction that comes from saving the releasing costs, which is the equivalent of decreasing the value of B. Thus, with this subject matter, the four-cell approach tends to reduce to two cells. Either approach to encouraging socially desired behavior could be used where a choice is present that does *not* involve a contingent event. In such a situation, one would concentrate on changing the values of A and B or of a, b, c, and d without one or more P's or probabilities to be concerned with.

28. This is the Hausman and Thaler study referred to in note 20 *supra*.

These statistical techniques can also decrease the gray area as to who will appear, which can be important in increasing the percentage of defendants released. This is so because many judges may be well aware that a high percentage of all defendants appear, say about 90 percent. They may, however, indicate that the problem is that they do not know in advance which released defendants would be likely to be in the 90 percent, and which released defendants in the 10 percent. More specifically, a judge might say he releases about 70 percent of all defendants because those 70 percent are the ones he feels are likely to appear. Another 5 percent he feels are very unlikely to appear. The middle 25 percent he is not sure about either way, but he tends to hold them rather than risk the embarrassment of having them not appear. If the size of this gray or unpredicted 25 percent can be reduced, that should mean a substantial increase in the percentage of defendants released.

The probability of a defendant deliberately choosing not to appear in court can also be reduced by more vigorously prosecuting those who fail to appear without an adequate justification. Increasing the probability of such prosecution and the invocation of a substantial penalty raises the expected value of appearing relative to not appearing for the released defendant, thereby increasing the likelihood that he will choose to appear. The possibility of forfeiting one's bond can also encourage appearance for those released on bond, rather than on their own recognizance.[29]

29. The problem of how to increase the probability of a defendant appearing in court can be thought of as a decision theory problem. The released defendant has basically two choices, either to appear or not to appear. He will appear if he perceives the expected value of appearing to be greater than the expected value of not appearing. The expected value of appearing equals the benefits of appearing (for example, getting one's bond back and avoiding prosecution) minus the costs of appearing (for example, being subjected to a trial or wasting time if there is no trial) with those benefits and costs discounted by the probability of their occurring. The expected value of not appearing equals the benefits of not appearing (for example, temporarily avoiding either a trial or waiting at the courthouse) minus the costs of not appearing (for example, losing one's bond and getting prosecuted for bond jumping) with those benefits and costs discounted by the probability of their occurring. Thus, defendants are more likely to appear if those perceived benefits and costs can be changed so that the expected value of appearing will be more often perceived as being the greater value. Notification systems have the effect of reminding defendants of the costs of not appearing and the benefits of appearing. Screening systems have the effect of finding defendants who are more likely to perceive the expected value of appearing to be greater than the expected value of not appearing.

3. MAKE MORE VISIBLE THE ERRORS AND COSTS OF HOLDING DEFENDANTS WHO WOULD APPEAR

The main reason judges are more sensitive to type 2 errors than type 1 errors is because the error of releasing someone who fails to appear is presently more visible than the error of holding someone who would appear. What needs to be done is to increase the visibility of the holding errors. One meaningful way to do that is to show for each judge serving in the same circuit what percentage of defendants he holds in jail prior to trial, and what percentage of his released defendants appear for their court dates. We would probably find a great deal of variation among the judges with regard to the percentage of defendants they hold. We would probably, however, not find so much variation among the judges with regard to the percentage of their released defendants who appear in court.[30]

If the judges were arranged on a list from the judge with the highest holding percentage to the judge with the lowest holding percentage, we might see that the judge holding the highest percentage of defendants (about 70 percent hold) has an appearance rate of about 95 percent. The judge holding the fewest defendants (about 20 percent) might have an appearance rate of about 90 percent. Thus, the highest hold judge could not justify his high holding rate on the grounds that he is getting a much better appearance rate than the judges with lower holding percentages. He also would find it difficult to argue that his sample of defendants is substantially different than the sample of defendants which his fellow judges have since there tends to be less shopping in arraignment proceedings than in trial proceedings. We are thus roughly indicating the number of errors of holding defendants who would appear through the use of these aggregate statistics, even though it is impossible to determine whether such an error has been made in an individual case where the defendant is held in jail prior to trial.[31]

30. Variation among the judges in holding rates is partly indicated by substantial variations in holding rates across cities. *See* W. THOMAS, BAIL REFORM IN AMERICA 37-64 (1976); NAGEL, *supra* note 16. Lower variation among judges on appearance rates and crime-committing rates is partly indicated by the lower variation in those rates across cities. *See* W. THOMAS, at 87-109, and NAGEL, *supra* note 16. The Thomas study showed that when felony holding rates dropped from 52 percent in 1962 to 33 percent in 1971, and misdemeanor holding rates dropped from 40 percent in 1962 to 28 percent in 1971, the appearance rates of the increasingly large number of defendants released only dropped from 94 percent to 91 percent.

31. If contrary to our predictions, we find that the low holding judges on a court have a substantially worse appearance rate (rather than about the same appearance rate) than the

The raw data for calculating hold percentages and appearance percentages for individual judges can be obtained from the docket sheets which are public information. Obtaining this data thus requires no special cooperation from the judges the way having them answer hypothetical sentencing cases would. If these lists were periodically made available, it is likely that judges who are holding a high percentage of defendants would have a tendency to move downward. They would be motivated partly because of (1) embarrassment in comparison to their fellow judges, (2) lack of appearance rate justification for their high holding rates, (3) a respect for the norm of uniformity among judges, and (4) in recognition of the high holding costs relative to releasing costs, to which we now turn.

In addition to publicizing the errors of holding defendants who would appear, the costs of such errors can also be publicized. Many judges may be unaware of how high the holding costs are, and how many different types of costs are involved. The most obvious is that of jail maintenance which may be quite substantial when one considers the length of time the average defendant is held, and the fixed and variable costs needed to provide for him during that time. The costs also include the lost gross national product which can be attributed to defendants being unable to earn or produce anything while they are in jail. That cost may also be substantial even if it is only figured at the minimum wage.

An additional cost that is hard to assess but should still be analyzed is the bitterness that is generated by being held in jail for a substantial time awaiting trial, and then having one's case dismissed or acquitted for lack of sufficient evidence of guilt. It might be interesting to know what percentage of defendants held in jail prior to trial do have their cases dismissed or acquitted, or receive sentences shorter than the time they have already served awaiting trial. Another cost that can be more easily assessed involves deter-

---

high holding judges, then revealing those differences should have the effect of encouraging the low holding judges to do more holding. Likewise, if contrary to our predictions, we find the judges have about the same holding rates (rather than substantial variation in their holding rates) and different appearance rates, then revealing those differences should stimulate the judges with lower appearance rates to be more selective in whom they release and possibly to raise their holding rates. If the judges have about the same holding rates and about the same appearance rates, that revelation is not likely to change any behavior, even though lowering their holding rates might save holding costs without a commensurate increase in releasing costs, and even though raising their holding rates might save releasing costs without a commensurate increase in holding costs.

mining the number of families who are on welfare as the result of a breadwinner being held in jail prior to trial. Another cost from unnecessarily holding the defendants prior to trial is the cost of jail riots due to overcrowding. In recent years in New York, Washington, D.C., and elsewhere, some jail riots have been substantially attributed to the overcrowding from pretrial detention.

Still another important holding cost is the cost to due process which stems from the increased probability of an innocent defendant being convicted because he was unable to adequately prepare his case while being held in jail, and because he made a substantially poorer impression on the judge or jury by being a pre-trial detainee. Research by the Vera Institute does show that pre-trial detention increases the probability of conviction even when all other relevant variables are held constant. Pre-trial detainees who may be innocent or whose guilt may be quite difficult to prove are also quite vulnerable to prosecution offers to reduce the recommended sentence to the time served awaiting trial if the defendant will plead guilty. Released defendants are not vulnerable to that particular kind of pressure. Pretrial detainees may also be more likely to be convicted regardless of their guilt because they get faster trials while the prosecution witnesses are still fresh. Some prosecutors may concentrate on prosecuting pretrial detainees to the neglect of released defendants out of a feeling that detainees should receive a priority or because speedy trial laws emphasize fast trials for detainees.

By publicizing for each judge the percent of defendants he holds and the appearance percent he attains, and by making more visible how much it costs to hold defendants in jail, we are in effect trying to make the ratio between each judge's individual holding costs and individual releasing costs approach the ratio between society's holding costs and society's releasing costs. As of now, the judge's individual costs seem to produce a much lower ratio than the social costs. In other words, as an individual, a judge stands to lose more by making a mistake by releasing a no show than by making a mistake of holding a defendant who would show, but society generally stands to lose more by a holding mistake than by a releasing mistake, given the relative social costs involved and the probabilities of their occurring.[32]

32. Some judges or legislators may have a more narrow or different concept of social costs than others have. For example, making more visible the holding cost of lost gross national

4. DECREASE THE COSTS OF RELEASING DEFENDANTS WHO FAIL TO APPEAR

There are two main costs involved in releasing a defendant who fails to appear. One is the cost of having to rearrest him. The other is the cost of any crime he might commit while released. That crime cost theoretically, however, is not supposed to be a major criterion in determining whether a defendant should be released, since regularly including it would amount to a constitutionally questionable system of preventive detention, as contrasted to including it in exceptional cases like those involving obviously dangerous defendants. The rearresting cost (and incidentally the crime-committing cost) can be reduced by having a better knowledge of where the defendant is through a system of pretrial parole. Such a system would involve releasing the defendant on the condition that he at least periodically report his whereabouts, or be subject to rearrest for failure to do so. Heavy supervision in all cases, however, might be more costly than the incremental rearresting and crime-committing savings over an unsupervised system.

A way to substantially reduce both the non-appearance and crime-committing costs would be to reduce the time from arrest to trial. The longer the defendant is out, the more the probability increases that he will negligently or deliberately fail to appear, and the more time he has to become involved in a new criminal act. The defendant's ultimate non-appearance is especially increased if during that intervening period he is repeatedly told to come to court which he does, and due to congestion and delay the court is unready to hear his case. Thus, the pretrial release problem is closely associated with the problem of reducing delay in the criminal justice system.[33]

---

product may be less impressive to some judges than other judges as compared to the extra jail maintenance costs even if both dollar amounts are equal. Likewise, some judges may be relatively unconcerned with jail maintenance costs if they are borne by the state government rather than by their own local government. This further illustrates the value of working with non-monetary measures which tend to express the relative perceived utility of alternative possibilities to the decision-makers rather than the more superficial, but often more measurable, monetary measures.

33. One possible benefit from delaying the trial of released defendants is that while on bond, they may feel constrained to be more law abiding in order to avoid having their bond and release revoked, and especially to avoid more severe treatment from the prosecutor and the court when the case eventually does come up. Release on bond in some cases, however, may encourage crimes that would not otherwise have occurred if the defendant engages in illegal behavior in order to pay a bondsman.

Delay can be reduced by having more judicial personnel, diverting more cases away from the criminal justice system or at least away from trial, and attempting to shorten the trial stage through pretrial discovery and random jury selection. Delay can also be reduced by better sequencing and scheduling of cases. For example, if all the short cases were given a priority, the average waiting time would be reduced, although a maximum time constraint should be placed on all cases. Theoretically by shortening the path from arrest to trial through the elimination of certain stages, delay could be substantially reduced, but doing so might violate federal or state constitutional constraints.

Releasing costs that relate to crimes committed by released defendants can be reduced by the same methods which can be suggested for crime reduction in general. They basically involve applying a model like that shown in Figure 6 to reduce the expected value of crime committing and increase the expected value of engaging in legitimate alternate activities. By analogy to Figure 6, these methods logically include increasing the probability of being arrested, convicted, and jailed which can come about through such means as a more efficient criminal justice system or through less due process. The expected value of crime committing is also reduced by decreasing the benefits of successful crime committing through such means as reducing the vulnerability of potential crime targets, and reducing the peer group recognition that criminals often receive by trying to redirect gang orientations. The expected value of crime is further reduced by increasing the costs of unsuccessful crime committing possibly through more severe punishments, or by having criminals suffer the opportunity costs of missed occupational opportunities by first providing them with some meaningful opportunities to miss.

One method that would increase the probability of appearance and also decrease the cost of rearresting defendants who fail to appear is the method of high bond setting. If the bond is high and sure to be forfeited for non-appearance, the defendant is more likely to appear. Likewise, if the bond is high, or at least high enough to cover the rearrest costs, then if the defendant fails to appear, his bond forfeiture can be used to reduce those rearrest costs to zero in the average case. This makes bond setting like a pollution tax which has often been proposed in environmental law (the tax on a given firm is to the total dollars desired to clean up the area as the firm's pollution is to the total pollution in the area). That kind of proportionate tax is designed to deter pollution, analogous to deterring non-appearance. If, however, a business firm cannot economically eliminate its pollution, then the tax it pays is used to clean up or reduce the damage its pollution has caused, analogous to paying the rearrest costs. This high bond approach has the undesirable effect of increasing the likelihood of holding defendants, unlike the other approaches which do not produce conflicting effects. This approach logically raises the question of how high the bond should be in order to increase the probability of appearance and decrease the releasing costs, without counter-productively increasing the occurrence of holding. That is the issue of devising optimum bond schedules to which we now turn.

III. THE BOND-SETTING DECISION ACROSS CASES

A. *Non-Discretionary Bond Schedules*

1. THE PROBLEM

Thus far, we have been discussing the releasing and bond-setting decision on an individual case-by-case basis although we used illustrative data representing an average defendant rather than a specific defendant. Now, however, we would like to discuss the problem of developing rules designed to cut across cases of a given type as is done in non-discretionary bond schedules which specify, for example, that bond for non-aggravated battery shall be $1,000.[34] Bond schedules like that are increasingly being used to set bond by the police when judges are not available at night or on weekends in misdemeanor cases. Bond that is determined by such a schedule can be subsequently lowered or raised in a judicial proceeding initiated by the defendant or the prosecutor. There may also be an increase in the use of such schedules by arraignment judges in view of their advocacy by various legal scholars.[35] At the turn of the century, judicial reformers advocated indeterminate sentences in hopes of obtaining enlightened individualized discretion in criminal cases.

34. *E.g.*, 110A ILL. REV. STAT. § 528c (1969) states: "Bail for misdemeanors (other than traffic or conservation offenses) punishable by fine or imprisonment in a penal institution other than a penitentiary, shall be $1,000."

35. M. Kelly, *Social Science Evaluation and Criminal Justice Policy-Making: The Case of Pre-Trial Release*, in PUBLIC POLICY EVALUATION (Dolbeare ed. 1975); Wisotsky, *Use of a Master Bond Schedule: Equal Justice Under Law?*, 24 U. MIAMI L. REV. 808 (1970); W. THOMAS, *supra* note 4 at 211-13 and 258-59; R. HAND & R. SINGER, SENTENCING COMPUTATIONS LAWS AND PRACTICE 2-39 (1974).

By mid-century, judicial reformers had become somewhat disillusioned with the feasibility of obtaining such enlightened discretion, especially in view of various studies which show how arbitrary sentencing and other forms of judicial discretion have become.[36] As a result, there is an increasing tendency to return to more objective although automatic standards as being preferable to arbitrary discretion and as being much more feasible than enlightened discretion.[37]

The problem here is one of scientifically determining what bond should be set for each crime and for different types of defendants where the characteristics of the defendant can be legally considered. It would be unconstitutional to have one bond schedule for blacks and one for whites, regardless of what might be shown scientifically about the relation between race and appearance rates. So long as the correlation is substantially short of a perfect correlation (for example, where no blacks appear and all whites do), we do not want to use such a correlation because it would result in too many type 2 errors of releasing white persons who should be held, and would especially result in too many type 1 errors of holding black persons who should be released. It would not, however, be unconstitutional to provide lower bonds for teenagers than for adults so as to facilitate the release of teenagers for whom the holding costs to society may be greater in terms of creating increased criminal behavior. A bond schedule that takes ability to pay into consideration would also probably be constitutional although it might be easier to administer such a schedule if ability to pay were dichotomized into indigent and non-indigent with indigency determined as part of the proceeding associated with appointing the public defender or assigned counsel. Just as bond schedules can vary from state to state,

36. M. Frankel, *supra* note 1; R. Dawson, Sentencing: The Decision as to Type, Length, and Conditions of Sentence (1969); A. Trebach, The Rationing of Justice: Constitutional Rights and the Criminal Process (1964); Nagel, *Disparities in Criminal Procedure*, 14 U.C.L.A. L. Rev. 1272 (1967); D. Fogel, ". . . We Are the Living Proof . . ." The Justice Model for Corrections (1975).

37. The Illinois legislature is currently considering the adoption of a sentencing system in which all sentences are fixed rather than indeterminate. The fixed terms are specified for certain types of crimes and defendants by statute with no judicial or parole board discretion, although discretion in the charging process is still available to the prosecutor. The fixed term can be reduced one day for each day of good time in prison, but the good time cannot be retroactively taken away if the defendant subsequently misbehaves. *See* J. Foster, Definite Sentencing: An Examination of Proposals in Four States (1976).

they can also probably vary by size of city within a state in view of the different relation between bond and the probability of appearing in rural, urban, and metropolitan areas.

2. THE GOAL AND A SOLUTION

The ideal or optimum bond for a given crime and type of defendant would be the bond that maximizes the probability of the defendant appearing in court (PA) while minimizing the probability of his being held in jail (PH). A major defect in this statement of the optimum is that it fails to consider that as the bond goes up, the probability of appearing goes up which is desirable, but the probability of being held in jail also goes up which is undesirable. Thus, we cannot simultaneously maximize PA while minimizing PH.

A better statement of the optimum bond for a given crime would thus be that the optimum non-discretionary bond is the bond that maximizes the difference between the probability of appearing and the probability of being held, or that maximizes PA — PH. For the average defendant in the data on which Figure 3 is based, the probability of appearance is .92, and the probability of being held in jail is .27. Thus, in the average case, the bond is being set at such a level that PA — PH equals .65. Perhaps through a scientifically determined bond schedule, that difference could be made even greater partly by raising PA, but mainly by lowering PH. Even if the bond schedule did not increase the .65 difference, it might still be an improvement over the prevailing system if it substantially decreases arbitrary and discriminatory bond-setting, particularly if there is a system for adjudicating cases where the defendant can show that the automatic bond is especially unreasonable in his specific case. Too easy a system for adjudicating the automatic bond, however, would lead to disparities that favor those who have access to expensive lawyers who could obtain a lower bond and disfavor those who do not.[38]

38. Usually if a defendant can afford an expensive lawyer, he can also afford a high bond. Thus, easy adjudication of the automatic bond would probably not favor the rich, but rather favor middle income defendants who are not rich enough to meet the high bond, but not poor enough to have to rely on the public defender. In other words, the rich defendant does not need an easy system for arguing that the automatic bond is unreasonable, and the poor defendant may have a lawyer who lacks the time and assistance to be able to effectively argue for a bond reduction. This kind of discriminatory pattern against the poor may, however, not be as great under automatic bond schedules as under traditional bond setting systems if the bond schedules tend to involve lower bonds in recognition of the fact that those bonds produce the best PA — PH values.

To set a bond figure that maximizes PA — PH, one could simply determine for a large set of past cases what was the average bond set for the disorderly conduct cases, the shoplifting cases, and so on. The past average would thus become the future automatic bond. This may be the way some bond schedules are established. That would be more scientific (that is, more based on empirical data) and possibly more rational (that is, more likely to maximize PA — PH) than a kind of gut reaction attempt to create a bond schedule. That average-bond method, however, presumes that past practice has been as rational as one can get. Such a presumption may be quite faulty.

A possibly more goal-effective alternative would be to take that large set of past disorderly conduct cases, shoplifting cases, and so on, and then group the cases for a given crime into various bond categories somewhat like those used in the questionnaire in Figure 6. The set of approximately ten categories used should vary with the crime, and they should be set up in such a way that each category has approximately the same number of cases.[39] For each bond category in the disorderly conduct cases, we would then determine what percentage of the defendants appeared in court (that is, PA for that category on that crime), and we would determine what percent of the defendants were held in jail (that is, PH for that category on that crime).

That information could easily be plotted on a figure like Figure 8. For each bond category, there is a dot corresponding to the percent of appearance for that category (PA), and an X corresponding to the percent of being held for that category (PH).[40] There are thus

39. Although somewhat more difficult to handle, the subsequent interpolation and curve-fitting would be more meaningful if each bond category involved an equal interval in the sense of covering an equal number of dollars rather than covering an equal number of cases. Equal-distance intervals (that is, equal dollars) are more difficult to work with than equal-frequency intervals (that is, equal cases) in the analysis which follows because the base of the percentages (that is, the number of cases) for each equal-distance interval may sometimes be too small to produce reliable percentages. A good compromise in the bond-setting context might involve seven or so equal-distance intervals or categories, with a zero category added saying "less than $1," and a last category added saying "more than D dollars." The "less than $1" category refers to release on recognizance, someone else's assurance of appearance, a specified cash bond which is not collected but which would be collected and forfeited if non-appearance occurs, or other release not involving a cash or property deposit. "D dollars" is the highest figure one can have and still have seven equal distance intervals between category zero and the last category, given the nature of the crime and the bond-setting data that relates to it.

40. Figure 8 is drawn rather roughly to illustrate the hypothesized general relations

ten pairs of dots and X's. To find the pair where PA — PH is greatest, all we have to do is just subtract each PH from its corresponding PA. Doing so will possibly tell us with this hypothetical data that bond category 4 involves the biggest difference between PA and PH. Therefore, we could specify that the midpoint of the interval for category 4 or the range for category 4 shall be the fixed bond for disorderly conduct in medium-size cities in the state of Illinois (that is, in cities between 100,000 and 1,000,000 population).[41] In other words, what we have done is in effect to treat the probability of appearing as being like income, and the probability of being held as being like expenses, and we are trying to find the optimum price or quantity to produce that will maximize income minus expenses.[42]

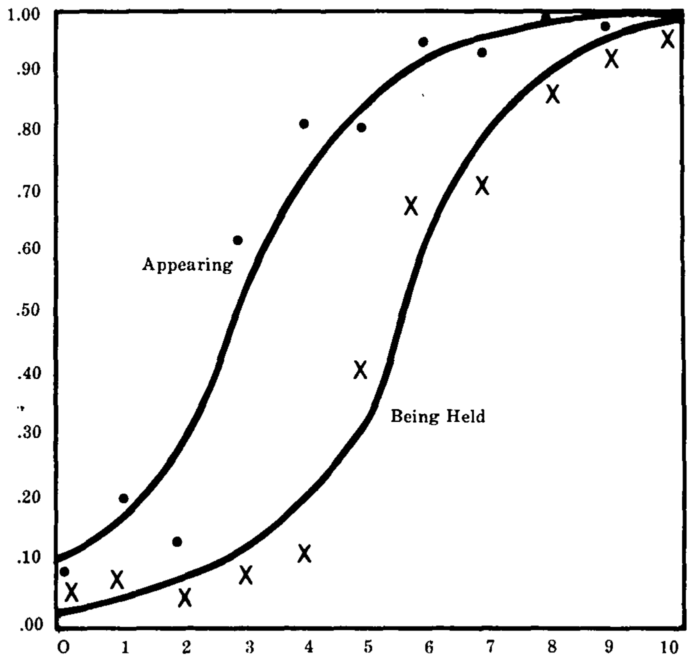
between PA and the bond level categories, and between PH and the bond level categories. A more exactly drawn PH curve would end at the zero point in the lower left hand corner of the graph if the zero level category meant release on one's own recognizance since there would then be zero probability of being held. Likewise, a more exactly drawn PA curve would curve upward at the zero bond level category because those who are released on their own recognizance usually have a high probability of appearing by virtue of their background characteristics, not because of the incentive to retrieve their bond. Thus, to make Figure 8 more meaningful for bond-setting purposes, the zero level or ROR category should probably not be included in the analysis.

41. This approach of trying to find a bond level for each crime that maximizes the probability of appearing minus the probability of being held can be contrasted with both the prevailing highly discretionary bond-setting and with non-discretionary schedules that merely codify previous average bond levels for given crimes or that use arbitrary bond levels to achieve objectivity but fail to adequately achieve pretrial release goals. By seeking bond levels that maximize PA minus PH, objectivity is likely to be increased. There is also likely to be a small increase or no change in appearance rates, and a substantial decrease in holding rates, since optimum bond levels in light of that PA — PH criteria are likely to be quite low, but without lowering the probability of appearing. A pretrial release system in which there are no money bonds at all might at first glance seem fairer in not discriminating against the poor. If, however, such a system allows considerable judicial discretion as to who will be held rather than released on one's own recognizance, then stereotypes and prejudices may result in more, not less, discrimination and more ineffectiveness in achieving the goal of maximizing appearance rates minus holding rates. A pretrial release system that completely abolishes money bonds is also not likely to be adopted.

42. As an alternative to Figure 8 where we plot PA and PH separately, we could plot one curve showing PA minus PH at each bond level category. Such a curve would tend to be relatively low at both the lowest and the highest bond level categories. It would be relatively high at a low to middling bond level category where in Figure 8 the gap between PA − PH is greatest. Such a hill-shaped curve could be adequately expressed by the equation $PA - PH = a + b_1C + b_2C^2$ where C represents the bond level category for say ten categories, and where the values of b_1 and b_2 are determined by a computerized regression analysis which involves providing the computer with ten sets of values for C (from 1 to 10), C squared (from 1 to 100), and the corresponding PA − PH from the data gathered. The optimum C (or C corresponding to the maximum PA − PH) is equal to $-b_1/2b_2$ by virtue of the algebra rules for finding the value of X where Y is a maximum, and Y has a quadratic relation with X.

FIGURE 8. THE RELATION BETWEEN LEVEL OF BOND AND BOTH THE PROBABILITY OF APPEARING IN COURT AND THE PROBABILITY OF BEING HELD IN JAIL

PROB. OF APPEARING (PA) & PROB. OF BEING HELD (PH)



BOND LEVEL CATEGORY

(Hypothetical data for the crime of disorderly conduct in medium sized cities in the state of Illinois)

3. VARIATIONS ON THE SOLUTION

With this data, however, we could try to fit a smooth curve to the PA dots and another smooth curve to the PH X's. We could then geometrically or algebraically determine where the greatest separation exists between those two curves. Doing so might reveal that the optimum bond category is 4.6, and by simple interpolation we could translate that into a dollar figure. The type of smooth curve that is likely to nicely fit our dots and X's is a third-degree polynomial or cube law or S-shaped curve of the form PA or PH = $a + b_1C + b_2C^2 + b_3C^3$ where C stands for category number. The a, b_1, b_2, and b_3 are coefficients whose value can easily be determined by feeding into a computer the ten PA's, the ten C numbers from 1 to 10, their squares, and their cubes, along with a linear regression program, and then doing the same thing for the ten PH's.[43] Once we have those numerical coefficients, then we want to find the slope of PA — PH relative to C. After we find that slope through the elementary rules for finding slopes, we want to set that slope equal to 0, and then solve for C in terms of the numerical coefficients. Where the slope of PA — PH relative to C becomes zero after rising up and before starting down is where PA — PH is at a maximum. Solving for C at that point involves solving a quadratic equation. Doing so and then interpolating will give us a more precise bond figure than merely finding the percent of appearance and corresponding percent of being held that have the biggest difference.[44] Finding a precise point within a bond category or interval (rather than simply using the range of the category or a range around the precise point) might, however, not be desirable because an optimum bond schedule should probably provide ranges for each crime which allow some discretion rather than a precise point or dollar figure which allows no discretion.[45]

If a legislature considers promoting appearance to be twice as important as promoting release, then the goal to maximize might be expressed as 2(PA) minus PH and the optimum bond interval

43. On fitting curves like third-degree polynomials to data like that shown in Figure 7, see H. BLALOCK, JR., SOCIAL STATISTICS 459-62 (1972); E. TUFTE, DATA ANALYSIS FOR POLITICS AND POLICY 108-34 (1974).

44. On finding the slope of a difference between two polynomials, setting the slope to zero, and then solving for the independent variable, see S. RICHMOND *supra* note 3, at 40-86; M. BRENNAN, PREFACE TO ECONOMETRICS 111-79 (1973). If the equation for the PA curve is $PA = a + b_1C + b_2C^2 + b_3C^3$, and the equation for the PH curve is $PH = a' + b_1'C + b_2'C^2 + b_3'C^3$, then $PA - PH = (a\text{-}a') + (b_1\text{-}b_1')C + (b_2\text{-}b_2')C^2 + (b_3\text{-}b_3')C^3$. That equation simplifies to $PA - PH = A + B_1C + B_2C^2 + B_3C^3$, where $A = a\text{-}a'$, $B_1 = b_1\text{-}b_1'$, $B_2 = b_2\text{-}b_2'$, and $B_3 = b_3\text{-}b_3'$. Given that equation, the slope of PA−PH relative to C is $B_1 + 2B_2C + 3B_3C^2$. Setting that slope equal to 0 and solving for the value of C involves solving the quadratic equation formula $C^{*2} = (-2B_2 \pm \sqrt{2B_2^2 - 12B_1B_3}) / 6B_3)$, since if $c + bX + aX^2 = 0$, then $X = (-b \pm \sqrt{b^2\text{-}4ac}) / 2a$.

45. The most meaningful way to relate bond level to either the probability of appearing or the probability of being held might be to plot each case individually rather than by groups, categories, or intervals which are subjectively arrived at. If each case is plotted individually, the dot plotted for PA would either be at 1.00 or at zero since the defendant would have either appeared or not appeared. With that set of data for a large number of cases, one can fit an S-shaped curve to the PA data by using the method described in note 44 *supra* and the accompanying text, except that each C would correspond not to a bond category number, but rather to the exact amount of the bond for each case. One can do likewise with the PH data. One can then find the exact bond level where PA — PH is at a maximum by using the method described in note 44 or 42 *supra* and the accompanying text. This more exact approach, however, may not be as simple to work with as the approach that involves grouped data, as described in note 39 *supra* and the accompanying text.

would be the one that is highest on that goal. In more general terms, one could say the goal to be maximized is W(PA) minus PH, where W is the desirability weight of PA relative to PH. The value of W could be less than one if PH is valued more highly than PA. For example, the value of W could be .5 if PH is considered twice as important as PA.

For a still more realistic way to express the goal for determining an optimum bond interval, one might use the expression $(PA)^W$ divided by PH. Doing so takes into consideration that as the probability of appearing increases, community satisfaction increases, but not at a constant rate in view of the fact that the first unit of a good thing normally produces more incremental satisfaction than the second unit. Likewise, the dividing takes into consideration that as the probability of holding increases, community satisfaction decreases, but not at a constant rate in view of the fact that the first unit of a bad thing normally produces more incremental dissatisfaction than the second unit. The W still represents the desirability weight of PA relative to PH, but it is now an exponent of PA rather than a multiplier. This goal is in effect like a weighted benefit-cost ratio to be maximized by finding the optimum or highest-scoring bond interval for each type of crime in each geographical area.

In addition to being useful for determining an optimum bond level for each type of crime, the kind of data shown in Figure 8 can also be helpful for obtaining a better understanding of the relation between bond setting and other variables. If the hypothetical data shown in Figure 8 is reasonably accurate, that means at the lower bond levels, the probability of being held is continuously quite low; and at the higher bond levels, the probability of being held is continuously quite high. Only in the middle levels is there a substantial variation between bond level and the probability of being held, or for that matter the probability of appearing in court. In other words, the average individual can equally meet a bond of $10, $20, $30, and so on at the low levels; and will equally fail to meet a bond of $100,000, $200,000, $300,000 and so on at the high levels. The probability of appearing may drop and stay especially low when the bond is low and a bond forfeiture is treated as a fine with little or no likelihood of prosecution for bond jumping. If our data is reasonably accurate, we can predict what PA and PH will be at various bond categories for the average disorderly conduct case in medium-size Illinois cities or whatever our data base is. That kind of predictive power could be helpful to judges in discretionary bond-setting cases either done geometrically with figures like Figure 8 or done algebraically with a third-degree polynomial equation. Thus, the kind of analysis represented by Figure 7 can be helpful in both the making of bond-setting statutes and in the resolving of specific bond-setting cases.[46]

B. *The Optimum Percentage to Release*

Another pre-trial release problem that cuts across cases is the problem of trying to decide approximately what should be the optimum percentage of defendants to release prior to trial. If too few are

46. S. A. Schaffer, Bail and Parole Jumping in Manhattan in 1967 (an unpublished paper of the New York Vera Institute of Justice) provides some empirical data on the relation between bond levels and both the probability of appearance and the probability of being held which reinforce the general nature of the hypothetical curves shown in Figure 8. With regard to PA, his data shows at page 29:

Bond Level	Prob. of Appearance	Bond Level	Prob. of Appearance
$ 1 - $ 25	.76	$ 251 - $ 500	.95
$ 26 - $ 50	.82	$ 501 - $1000	.93
$ 51 - $100	.81	$1001 - $2500	.89
$101 - $250	.84	Over $2500	(1.00)

Probabilities in parentheses indicate that the number of cases was between 11 and 50, whereas probabilities not in parentheses indicate the number of cases was over 50. The general pattern shown here is that the probability of appearance does generally increase as the bond level increases although that pattern may have been clearer if separate charts were shown for each crime rather than combining all cases together. One would expect higher bonds for more severe crimes and generally a lower probability of appearance for more severe crimes, which means crime severity should be held constant in order to determine the relation between the bond level and probability of appearance.

With regard to PH, Schaffer's data shows at page 23:

Bond Level	Prob. of Being Held	Bond Level	Prob. of Being Held
Under $100	.03	$1001 - $2500	.80
$101 - $500	.57	Over $2500	.96
$501 - $1000	.65		

The general pattern shown here is that the probability of being held does clearly increase as the bond level increases. That is probably true regardless of the crime although some crimes are disproportionately committed by poorer people who are less able to meet a given bond level. Note that (as in Figure 8) PA tends to be above PH for a given bond level. They come close together at very high bond levels, and seem to be farthest apart in the second category of $26 to $50. The fact that Schaffer uses different bond level categories for the two sets of relations complicates the analysis and illustrates the subjectivity of arriving at appropriate bond level categories as mentioned in notes 45 and 39 *supra*.

released, there will be excessive holding costs although releasing costs will be down. If too many defendants are released, there will be excessive releasing costs although holding costs will be down. An optimum percentage to release can be found in a manner similar to the Figure 7 approach to finding an optimum bond level. We can in effect plot total holding costs and total releasing costs for different releasing percentages in the same city at different points in time, in different cities at the same point in time, or a combination of both. The optimum percentage is the percentage where the sum of the holding costs plus the releasing costs is at a minimum, or where the sum of the holding benefits plus the releasing benefits is at a maximum.

That kind of analysis is the subject of a separate paper.[47] What we briefly want to discuss here is how that kind of analysis relates back to pretrial release decisions in individual cases and in nondiscretionary bond schedules. All three kinds of analysis will be reconcilable if they all consider the same values in the same cases. Thus, when that occurs, if the optimum percentage analysis indicates that total costs are minimized when 96 percent of the defendants are released, then the bond schedule approach will also result in releasing 96 percent of the defendants as will the individual case approach. We cannot, however, be sure that the judges in the nonmonetary individual case analysis are operating under the same values as we included in the monetary case analysis. We know that in our bond schedule analysis, we did not use the same values as we used in the monetary individual case analysis because of the need to simplify when one is working across cases rather than in an individual case. We further know that the optimum percentage to hold is not the same as the actual percentage held.[48]

These variations inform us that our models are not completely capturing empirical reality with regard to what goes on in pretrial

47. S. NAGEL, *supra* note 16.

48. The 96 percent optimum release level arrived at in S. NAGEL, *supra* note 16 and the optimum bond amount of $26 to $50 arrived at in note 46 *supra* would be equal to each other in their results if they were based on the same empirical data and the same normative values. The 96 percent optimum release figure, however, is based on data from a sample of 23 cities, whereas the $26 to $50 optimum bond figure is based on data from New York City, although both figures are for approximately the year 1969. Likewise, the 96 percent optimum release figure is based on more considerations than just maximizing PA — PH. Those additional considerations include minimizing pretrial crime-committing and having monetary weights associated with PA and PH.

release. Those discrepancies could possibly be remedied by gathering more questionnaire and other data in order to make our models fit reality better. Those discrepancies could also possibly be remedied by trying to get judges and legislators to move closer to the models if the models have some aspects that are more rational and objective than may be present in the prevailing decision-making. Perhaps what is needed is more realistic social science and more rationalistic governmental decision-making, both at the same time.

IV. Some Tentative Conclusions

In addition to the above general statement on the need for realism and rationalism, it seems appropriate to also conclude by stating what specific hypotheses our deductive analysis has implicitly generated that might now be explicitly tested with additional data. One set of hypotheses are largely methodological in nature and relate to the measuring tools which have been presented. We in effect hypothesized that the questionnaire approach included in Figure 2 would be a meaningful way to determine the satisfaction and dissatisfaction received by an arraignment judge from releasing a defendant who would or would not appear or from holding a defendant who would or would not appear. The same kind of measuring instrument involving rank ordering, anchoring, and then relative numerical measures could also be applied to other decision makers in the criminal justice process. The second measuring tool presented involved multiple sets of hypothetical situations like the set included in Figure 6. We implicitly hypothesized that such an approach would be a meaningful way of determining the relative importance of various criteria in bond-setting and other decisions more so than an approach directly asking decision makers what criteria they use. Likewise, the decision theory questionnaire of Figure 2 seems to be a more meaningful way of determining threshold probabilities for releasing defendants than directly asking a judge what his threshold probability is.

In addition to the measuring instruments, our analysis has implicitly stated a number of substantive hypotheses that might merit further testing. They include such statements as: (1) some arraignment judges are mainly oriented toward avoiding the release of bad-risk defendants and a probably smaller group of other judges are mainly oriented toward avoiding the holding of good-risk defendants; (2) there are substantial relationships between the background and attitudinal characteristics of arraignment judges and their orientations toward avoiding type 1 and type 2 errors; (3) there are substantial relationships between the characteristics of the defendants and their crimes and the threshold probabilities indicated by the responding judges to the decision theory situations; (4) there are substantial relationships between the characteristics of the cities, particularly regarding holding and releasing costs, and the threshold probabilities indicated by the responding judges; (5) judges who tend to perceive the probability of appearing as being low are judges who demand a high probability of appearing before they will release; (6) in the five above hypotheses, one could substitute probability of not committing a crime for the probability of appearing in order to test the role of both probabilities in pretrial release decisions; (7) arraignment judges tend to set bond in terms of past average bonds for certain crimes rather than in terms of an analysis of the defendant's probability of appearing in court at various bond levels; and (8) the probability of appearing and the probability of being held in jail bear a positively-sloped S-shaped relationship with bond levels like that shown in Figure 7.

In addition to the above substantive hypotheses which emphasize understanding why variations occur in pretrial release, the analysis also generates hypotheses concerning how to improve the pretrial release system in light of given goals. For example: (1) if judges are informed how their threshold probabilities, holding percentages, and appearance percentages compare to their fellow judges, then those judges who are relatively more different will tend to change their attitudes and behavior more toward the average; (2) if data is obtained on holding costs and releasing costs from various cities, one will find that the expected value or cost of releasing is substantially less than the expected value or cost of holding for the average defendant, thus supporting the rationality of releasing the average defendant although the costs may exceed the benefits for certain types of defendants; (3) if arraignment judges want to maximize the benefits minus the costs in a bond setting decision, they should decide in terms of an analysis of the defendant's probability of appearing in court at various bond levels; (4) if legislators, state supreme courts, or other bond-schedule makers want to maximize the probability of appearing minus the probability of being held, they should adopt the bond level for each crime where (in the past) the separation between those two probabilities or percentages has been greatest; (5) if arraignment judges want to minimize the sum of the holding costs and releasing costs, they should hold only about five percent of all defendants; and (6) to more accurately determine the probability of a defendant appearing in court, an arraignment judge should make some use of the multiple-variable prediction schemes developed by such socio-legal programs as the Vera Institute.

In light of the above implications raised by the analysis, one can possibly reach the overall conclusion that decision theory is a useful approach for generating some insights, hypotheses, and explanations with regard to pretrial release decisions. One might also be able to see how decision theory can be a similarly useful approach with regard to other decisions in the criminal justice process that relate to contingent events, such as the decision by a police officer to arrest rather than issue a summons, a prosecutor or defense counsel to go to trial rather than accept an out-of-court settlement, a sentencing judge or parole board to incarcerate or continue incarceration, a jury to find liability or convict, and a lawyer to appeal. Decision theory is a provocative way to conceptualize what is and what ought to be involved in decision-making. The approach in itself does not provide hard data answers, but it may serve a useful function by providing many questions and also by providing a means for integrating the answers being developed.[49]

49. The measuring instruments, the substantive or causal hypotheses, and the means-ends or prescriptive hypotheses mentioned above are the subject of an empirical research design proposal which has been submitted by Nagel and Neef to the National Science Foundation and other funding agencies under the title "Decision Theory and the Criminal Justice System." The research design basically involves working with 20 federal and state court systems to (1) increase the sensitivity of criminal justice decision-makers to avoiding type 1 errors, (2) develop more objective and effective decision-making guidelines for criminal justice decision-makers, and (3) to better understand the threshold probabilities of criminal justice decision-makers. A copy of the proposal is available from the authors on request.

V APPENDICES

A. *Glossary of Symbols*

SYMBOL	REPRESENTS	FIRST APPEARING
ERRORS, TYPE OF:		
Type 1 error	An error of holding a defendant who would appear or who should be released (rejecting a true hypothesis or presumption)	I-A1
Type 2 error	An error of releasing a defendant who would fail to appear or who should be held (accepting a false hypothesis or presumption)	I-A1
EXPECTED VALUES:		
EV_H	Expected value of holding a defendant	II-B
EV_R	Expected value of releasing a defendant	II-B
HOLDING COSTS AND BENEFITS:		
HC	Holding costs in keeping a defendant in jail pending trial	I-A3
HB_1	Holding benefits which result from rearresting releasing costs being saved when defendant is held	I-A3
HB_2	Holding benefits which result from crime-committing releasing costs being saved when defendant is held	I-A3
JUDGES, TYPE OF:		
Judge 1A	Judge who is more worried about holding a good-risk defendant than releasing a bad-risk defendant	I-A1
Judge 1B	Judge who is more worried about releasing a bad-risk defendant than holding a good-risk defendant	I-A1

PROBABILITIES, SEPARATE:

PA, or just P	Probability that defendant will appear or percent of defendants who have appeared for a given bond-setting category	I-A1
PA*, or P*	Threshold probability of appearance that has to be met before judge will release a defendant in a given situation	I-A1
PC	Probability or percent of defendants committing a serious crime while released	I-A3
PF, or 1-P	Probability or percent of defendants failing to appear in court	I-A3
PH	Probability of defendant being held in jail or percent of defendants that are held for a given bond setting category	II-B
PN	Probability of defendant not committing a serious crime if released	I-A2
PN*	Threshold probability of defendant not committing a crime that has to be met before judge will release a defendant in a given situation	

PROBALITIES, COMBINATION:

A	Probability of appearing in court and not committing a crime while released	I-A3
B	Probability of failing to appear in court and not committing a crime while released	I-A3
C	Probability of appearing in court and committing a crime when released	I-A3
D	Probability of failing to appear in court and committing a crime while released	I-A3

RELATING PA AND PH TO BOND LEVEL:

a	Value of PA or PH when bond-setting category C or level equals zero	II-B
b_1, b_2, b_3	Ratio between a change in PA or PH and a change in C, C^2, or C^3	II-B
C	Bond-setting category number from 1 to 10 that was used in specific cases to be fed into a computerized regression analysis to obtain values for a, b_1, b_2, and b_3	II-B

RELEASING COSTS AND BENEFITS:

RB	Releasing benefits, or the holding costs saved by releasing a defendant	I-A3
RC_1	Releasing cost of rearresting an average defendant, estimated at $200 each	I-A3
RC_2	Releasing cost of a crime committed by an average defendant while released, estimated at $1,000 each	I-A3

SATISFACTION ASSOCIATED WITH DIFFERENT OCCURRENCES:

Cell a, or —B	Relative dissatisfaction received by a judge if defendant fails to appear after being released	I-A1
Cell b, or +B	Relative satisfaction received by a judge if defendant appears after being released	I-A1
Cell c, or —A	Relative satisfaction received by a judge if he holds a defendant when he would have failed to appear if released	I-A1
Cell d, or +A	Relative dissatisfaction received by a judge if he holds a defendant when he would have appeared if released	I-A1

MISCELLANEOUS SYMBOLS:

ROR	Release of the defendant on his own recognizance without any bond	I-A1

B. *Basic Formulas Used*

Note: See the glossary for the definition of the symbols in the context of the formula presented.

(1) Expected value of releasing, non-monetary values
$EV_R = a(1\text{-}PA) + b(PA)$

(2) Expected value of holding, non-monetary values
$EV_H = c(1\text{-}PA) + d(PA)$

(3) Threshold probability for releasing, non-monetary values
$PA^* = (a\text{-}c) / (a\text{-}b\text{-}c+d)$
$PA^* = B/(A+B)$ in simplified version
$PA^* = 1/(X+1)$ in a further simplified version where $X = A/B$

(4) Expected value of releasing, monetary values
$EV_R = (RB+RC)(1\text{-}PA) + (RB+RC)(PA)$
where RC is a negative number, and there are no releasing costs when the defendant appears

(5) Expected value of holding, monetary values
$EV_H = (HB+HC)(1\text{-}PA) + (HB+HC)(PA)$
where HC is a negative number, and there are no holding benefits when the defendant would have appeared

(6) Threshold probability for releasing, monetary values
$PA^* = [(RB+RC) - (HB+HC)] / RC\text{-}HB$
where a negative number divided by a negative number is a positive number, which as a probability should be between 0 and 1, unless releasing (or holding) always produces more net benefits (B-C) than holding (or releasing) regardless of PA

(7) Decision rule on releasing or holding
Release if $EV_R > EV_H$ (*i.e.*, if $PA > PA^*$)
Hold if $EV_H > EV_R$ (*i.e.*, if $PA^* > PA$)

(8) Combined probabilities
$A = (PA)(PN)$ $B = (PF)(PN)$
$C = (PA)(PC)$ $D = (PF)(PC)$

(9) Relating PA and PH to bond level category

$$PA = a + b_1 C + b_2 C^2 + b_3 C^3$$
$$PH = a' + b'_1 C + b'_2 C^2 + b'_3 C^3$$

(10) Optimum bond-level category

$C^* = C$ where PA - PH is a maximum positive difference

$$C^* = \left(-2B_2 \pm \sqrt{2B_2^{\ 2} - 12B_1 B_3}\right) / 6B_3$$

(11) Probability of appearing for all defendants

$PA' = (1 - PH)(PA) + (PH)(1 - X \cdot PF)$

where PA and PF apply to released defendants, and X equals how many times greater than PF is the probability of a held defendant failing to appear if he were released