DETERRENCE AND SUBJECTIVE PROBABILITIES OF ARREST: MODELING INDIVIDUAL DECISIONS TO DRINK AND DRIVE IN SWEDEN

PERRY SHAPIRO*
HAROLD L. VOTEY, JR.

Illegal behavior responding to subjective probabilities of arrest and consequent costs is modeled within a framework of individual choice under uncertainty. The model is formulated for the case of driving while intoxicated and tested with data that include all arrests for drunken driving in Sweden from 1976-1979. Results indicate that an arrest experience reduces the probability that a person will drive while drunk. The results suggest that an arrest increases a person's perceived probability of arrest and/or the unpleasantness of an arrest and thus leads to a reduced chance of acting illegally.

I. INTRODUCTION

Deterrence theory is a theory about how people perceive and respond to the likelihood of threatened punishment. Yet though deterrence theory is both individual and subjective, much of the research that purports to test this theory investigates aggregate responses to objective measures of likely sanctioning. The use of objective measures aggregated over individuals does not signify fundamental disagreement over the nature of deterrence theory but instead reflects the nature of the available data. Governments routinely report information

* This paper has benefited from presentations at the Annual Meetings of the Law and Society Association, Denver (1983), the Crime Control Theory Conference, Carnegie-Mellon University (1983), and the Law and Society Workshop, U.C. Berkeley. We are particularly indebted to Richard Berk, Jacqueline Cohen, Samuel L. Myers, Jr., Llad Phillips, and Daniel Rubinfeld for constructive comments and to the editor and referees for extraordinary help in the development of this paper. The research has been made possible by support from the (Swedish) National Council for Crime Prevention, the (U.S.) National Institute of Justice, and the National Science Foundation's Law and Social Science Program. In addition, a large number of individuals and organizations have helped in the acquisition of the Swedish data. Of particular help in that connection have been Göran Hedström, Torbjörn Israelson, Eckart Kühlhorn, Ernst Nilsson, Peter Simon, and Henrik Thom.

on crime rates and sanctioning activity, but they cannot provide measures of how individuals regard the probability and likely severity of the punishments that attach to deviant behavior.[1] A major purpose of this article is to show that using the kind of objective information that governments often collect and with a minimum of additional plausible assumptions, one can model subjective probabilities of the certainty and severity of punishment and the role that these play in individual responses to law enforcement measures. The behavior we model is driving after drinking and the way it responds to official efforts at control. Thus, our second goal is to say something about this issue. Our findings are necessarily tentative because we have data on only a few of the characteristics that, in theory, have a place in our model. However, what we do find appears entirely plausible, given what we know about deterrence, learning behavior, and the kinds of people who drink and drive.

It appears from the literature on drinking after driving that innovative law enforcement techniques and newly increased sanctions, even when coupled with educational efforts indicating the new risks, seldom if ever have any lasting impact on the tendency to drink and drive.[2] However, virtually all methodologies associated with existing studies have been subjected to criticism. Studies of specific interventions using interrupted time series techniques often do not adequately control for alternative forces influencing accidents before, during, and following specific interventions.[3] Simultaneous econometric studies that more completely model accident generation are subject to the various criticisms that may be leveled against simultaneous studies in general.[4] And all this research shares the problem we have alluded to; it uses aggregate data and thus cannot identify the underlying behavior in which deterrent effects are rooted.

Individual decision-making is a personalized process, and drunken driving like other criminal behavior reflects individual characteristics. We know, for example, that offense and accident patterns vary by age and sex and that a large fraction of both accident-involved and convicted drunken drivers are

---

[1] Some interview studies do measure subjective probabilities, but they are typically based on small or biased samples, and the behavior they study is often not criminal or, if criminal, not very serious.

[2] A balanced survey, then current, was Jones and Joscelyn (1978). A more recent discussion is Ross (1981).

[3] Votey (1984) indicates the nature of the problem. Phillips *et al.* (1984) demonstrates that the technique's faults can be remedied by careful modeling.

[4] The most complete discussion is in Blumstein *et al.* (1978).

people with alcohol problems. It is also generally believed that individuals with prior convictions are more likely to be arrested and convicted of drunken driving than individuals without prior arrest records. From this knowledge it is reasonable to infer that individuals arrested with prior convictions are more likely than first-time offenders to be heavily involved with alcohol.[5]

As we shall show, it is possible to sort out effects related to prior violations as distinct from age effects, enforcement intensity, and sanction strength. In doing so, we can also address such issues as the effects that different types of sanctions have on differently situated individuals and the implications of generalized moral attachments to the law for drunken driving behavior (cf. Norström, 1981). However, to separate these effects, even with detailed data, requires a carefully worked out evaluation strategy and careful hypothesis testing. A behavioral model must be constructed that permits a test of the deterrence hypothesis within a framework accounting for the multidimensionality of individual differences. It must also take account of the fact that the actual level of drunken driving behavior is not known. Our task begins with the theoretical development of the model.

II. MODEL OF BEHAVIOR[6]

This study treats individuals as decision-makers who act on their own volition. Following in the tradition of Bentham (1948) and Beccaria (1963), we assume that individual behavior is governed by perceptions of self-interest. If these assumptions of rationality are wrong, our model is misguided and cannot be expected to fit the data.

Thus, in attempting to model the factors that are conducive to or inhibit driving after drinking, we are implicitly claiming that people in some sense calculate the costs and rewards of this behavior. It may be before the drinking begins, as when people arrange to have a sober companion available to drive them home, or it may be while "under the influence," as when people call taxis to drive them home or at least allow others to call taxis for them. If our model fits the data, it corroborates

[5] See, for example, articles by I. Sandler *et al.*, P.F. Zelhert, Jr. *et al.*, M.L. Selzar, and R.B. Voas in Israelstam and Lambert (1974).

[6] A similar model using subjective probability was developed by Shapiro (1980) to study income tax evasion. Tax evasion is a crime similar to drunken driving in that the overall level of the activity is unknown. The tax evasion model is presented in an Internal Revenue Service document (1980).

the work of those who have looked at specific interventions and who conclude that relatively little if any drunken driving is done by people so crazed by alcohol or oblivious to its effects that they cannot appreciate the fact that it may be in their interest and the interest of others not to drive after drinking.

We treat the individual as a person confronted with four choices:

1) to refrain from both drinking and driving;
2) to drive but refrain from drinking;
3) to drink but refrain from driving;
4) to both drink and drive.

A person who chooses any of the first three possibilities will not be arrested for drunken driving. One who selects the fourth choice will drive while intoxicated and may or may not get caught. The self-interestedly rational individual whom we are modeling will evaluate the consequences of the various outcomes and act to maximize expected utility.

The utility of behaving lawfully is

$$U(L) = \max\{U(1), U(2), U(3)\}, \qquad (1)$$

where $U(1)$, $U(2)$, and $U(3)$ are the utilities of choices, $i = 1, 2$, or 3. In other words, a person can choose from among three types of lawful behavior and will choose the option that is most satisfying. While there is, inescapably, some uncertainty about how satisfying a choice will turn out to be, the expected utility of the lawful choices does not, in some obvious way, turn on contingencies largely outside the actor's control. The situation is quite different with respect to choice 4, the decision to drink and drive. Here the decision-maker may anticipate either $U(N)$, the utility of breaking the law and not being arrested, or $U(AR)$, the utility of breaking the law followed by an arrest. Thus, the expected utility of driving after drinking ($U(DD)$) is a function of the utility of the two possible outcomes weighted by the perceived probabilities that they will be realized. In other words, $U(DD)$ depends on what people *believe* are the costs of an arrest (i.e., the pain they associate with fines, jail, loss of license, stigma, etc.) and on their perceptions of the probability they will be arrested. We may represent this formally as $U(DD) = \tilde{P}(AR)U(AR) + (1-\tilde{P}(AR))\ U(N)$, where $\tilde{P}$ denotes a subjective rather than an objective probability. What this means is that the expected utility of driving after drinking equals the subjective probability[7] that

[7] Our use of subjective probabilities is common to most models of decision-making under conditions of uncertainty and is consistent with deterrence theory. The idea of a subjective probability recognizes the fact that people often don't know what the "true" probabilities of an event are and

one will be arrested $[\tilde{P}(AR)]$ times the pain of arrest plus the subjective probability that one will escape arrest times the "pleasure" of driving while intoxicated.[8]

A person chooses to drink and drive, given our assumptions about rational behavior, if the expected utility from doing so exceeds the utility from behaving lawfully. In symbols, the decision is to drink and drive if, and only if,

$$\tilde{P}(AR)U(AR) + (1 - \tilde{P}(AR))U(N) \quad U(L). \tag{2}$$

Manipulation of this inequality implies that an individual will drink and drive if, and only if,

$$\tilde{P}(AR) < \frac{U(N) - U(L)}{U(N) - U(AR)}. \tag{3}$$

This means that drunken driving can be expected only when the subjective probability of arrest is less than the difference between the utility of successful drunken driving (i.e., drunken driving that does not end in arrest) and lawful behavior, divided by the difference between the utility of successful drunken driving and arrest. We shall call the numerator the marginal utility of successful drunken driving over lawful behavior $[MU(N,L)]$ and the denominator the marginal utility of successful drunken driving over arrest $[MU(N,AR)]$. The reasonableness of this result is evident from the various relationships it implies. First, the probability of drunken driving diminishes as the expected probability of arrest or pain of arrest increases. Second, the probability of drunken driving increases as the utility of lawful behavior diminishes. If lawful behavior is even more painful than the expected disutility of arrest, drunken driving will always occur. Third, the probability of drunken driving increases as the utility of successful drunken driving increases. Finally, the probability of drunken driving is lower, the more satisfying lawful behavior is relative to arrest. These propositions are all consistent with common sense.

Next, the model assumes that personal characteristics determine for each individual the ratio $MU(N,L)/MU(N,AR)$,

allows for tendencies to persist in misestimations of probabilities despite strong evidence of error. One virtue of the model we shall develop here is that it allows us to determine whether subjective probabilities differ systematically from the objective ones.

8 We assume the utility of an arrest is negative and shall at times refer to negative utility as "pain." The pleasure of driving after drinking is, no doubt, largely a matter of convenience, such as getting home quickly and having one's car at home in the morning. To the extent that there is sheer joy in driving while intoxicated, $U(AR)$ should be discounted by the driving that occurs before arrest, but we assume such pleasure is trivial compared to the costs of arrest and so ignore it in our model.

which is the right-hand side of the inequality in Equation (3). We might express this condition as:

$$\frac{MU(N,L|\varkappa)}{MU(N,AR|\varkappa)} = \frac{\Gamma(\varkappa)}{\tilde{C}(\hat{\varkappa})} \cdot V, \tag{4}$$

where $\varkappa$ and $\hat{\varkappa}$ are vectors of personal characteristics such as age, sex, a history of alcoholism, attitudes toward temperance, and past experience with the law. This expression posits a distinct set of characteristics, $\varkappa$, that affects attitudes and another set, $\hat{\varkappa}$, that affects how costs are perceived. We cannot, however, with our present data, identify which characteristics affect attitudes and which affect cost.

The $\Gamma()$ function is intended to capture underlying attitudes towards drinking and the obligation to obey the law. These might be affected, for example, by a person's age, sex, religion, and history of alcoholism. The $\tilde{C}$ () function reflects a person's estimate of the cost of arrest. This subjective estimate might be affected by sanction structures and past experience. A person who has been arrested might rate the cost of an arrest higher or lower than someone who has not had this experience. $\tilde{C}$ () is in the denominator of the expression because we can expect that the higher the subjective costs of arrest, the greater is the disutility of arrest and hence the larger $MU(N,AR|\varkappa)$. In other words, as the perceived costs of an arrest go up, the attraction of drunken driving will diminish even if the rewards of successful drunken driving retain their advantage over those of lawful behavior. This is another common-sense proposition. V is a random error term that we must take account of when we estimate the model because we cannot observe every relevant individual characteristic.[9]

From the foregoing discussion and Equations (3) and (4), we can define the probability that a person with certain attitudes and characteristics will drive after drinking as:[10]

$$P(DD|\varkappa, \hat{\varkappa}) = \frac{\Gamma(\varkappa)}{\tilde{P}(AR)\tilde{C}(\hat{\varkappa})}. \tag{5}$$

---

[9] In fact, our data set includes only a few individual characteristics. We assume that the error this leads to is uncorrelated with the variables that make up the vectors $\varkappa$ and $\hat{\varkappa}$.

[10] Letting lower case letters indicate the logarithm of the corresponding upper case variable and combining (3) and (4), the probability that a person will drink and drive is the probability that

$$v > \tilde{p}(AR) + \tilde{c}(\hat{\varkappa}) - \gamma(\varkappa).$$

If v is distributed exponentially, the probability that an individual drinks and drives (DD), conditional on $\varkappa$ and $\hat{\varkappa}$, is

$$P(DD|\varkappa, \hat{\varkappa}) = exp\{\gamma(\varkappa) - \tilde{p}(AR) - \tilde{c}(\hat{\varkappa})\}.$$

The equation in the text follows from the definitions of the variables.

https://doi.org/10.2307/3053448 Published online by Cambridge University Press

Or, in words, the likelihood that a person will drive after drinking will be less (a) the greater the person's allegiance to the legal system, (b) the less the person values driving after drinking, (c) the greater the perceived probability of arrest, and (d) the greater the perceived costs associated with arrest.

The relevant decision variables, are, unfortunately, impossible to measure given the available data. To take account of them, we model the relationship between subjective costs and probabilities and their objective counterparts.

In the case of arrest, the following relationship between subjective and actual probabilities is proposed:

$$\tilde{P}(AR) = P(y)^{\alpha(\varkappa)}, \tag{6}$$

where P is the actual probability of arrest. It is written as a function of environmental variables y (as contrasted with personal characteristics $\varkappa$). The parameter α, which is a measure of the accuracy with which arrest probabilities are perceived, is a function of $\varkappa$ because we believe it differs for different individuals. For example, a previous arrest may be associated with a person's probability assessment in two ways. The first might be called a learning effect. An arrest is likely to provide a person with new information about the criminal justice system, including, perhaps, better information about the true chance of an arrest. People who break the law and are not arrested may lower their estimates of the probability of arrest, and the experience of being arrested may lead to an upward revision in their subjective probability assessment.[11] Everything else being equal, including past criminal behavior, we would expect a person with an arrest record to see an arrest as more probable than one who had never been caught.

The second possibility is a form of selection bias. People may be arrested in part because of their carelessness, and this may reflect an unduly low assessment of the probability of arrest in the first instance. If this is not offset by the experience of arrest, one would expect arrest records to characterize those who underestimate the probability of arrest.[12]

Thus, Equation (6) tells us that the actual probability of arrest is determined by environmental factors (e.g., the density of police patrol), and the subjective probability of arrest is determined by this actual probability as inflated or deflated by the implications of past arrests. Equation (6) does not, by itself,

[11] The revision may be a discontinuous jump in the prior.

[12] Alternatively, this group may have an atypically high taste for risk.

allow us to separate learning effects from selection bias, but it does allow us to estimate their net effects. The value of the exponent, α, indicates which effect dominates. Because the true probability, P, is always between 0 and 1, a value of α larger than 1 indicates that a person *under* estimates arrest probabilities; an α equal to 1 indicates an accurate estimate, and an α smaller than 1 indicates an *over* estimate.

There may be similar learning effects and selection biases confounding the interpretation of subjective costs. People may act in ways that lead to single or repeated arrests because they associate relatively few costs with the experience of arrest. In addition, the special costs of early arrests in terms of stigma and the like may be such that people rationally see the costs of early arrests as higher than the costs of later ones. On the other hand, people may not appreciate the costs of an arrest until they have one. Or the threat of escalated penalties may lead people to fear later arrests more than earlier ones. For these reasons our specification of the relationship between subjective and objective costs is like our specification of the relationship between the subjective and objective probabilities of arrest in that it allows for individual variation. It is

$$\tilde{C}(\hat{\varkappa}) = \beta(\hat{\varkappa}, C), \qquad (7)$$

where the value of the measured costs, C, depends on the sanctions imposed.

The probability of arrest, given certain observed characteristics, is simply the probability that a person with those characteristics will drive while intoxicated times the probability that the person will be arrested while doing so. That is, using Equations (5), (6), and (7), we can specify

$$P(AR|\varkappa,\hat{\varkappa}) = \frac{\Gamma(\varkappa)}{\tilde{C}(\hat{\varkappa})} P^{1-\alpha(\varkappa)}. \qquad (8)$$

Equation (8) is essential to what follows and it is important that it be understood. It is a straightforward application of Bayes Law with a crucial assumption implied, namely, that actual probability of arrest does not depend on individual characteristics. This condition is violated if, for instance, the police have previously identified and are, thus, more likely to arrest alcoholics than they are non-alcoholics who drive while intoxicated.[13]

Equation (8) says that the probability that a person with characteristics $\varkappa$ (let us call this person A) will be arrested for

[13] We owe this observation to Llad Phillips. The question of who the police choose to arrest is interesting and important, but it is beyond the scope of this paper, and it cannot be answered with our data.

drunken driving varies proportionately with the benefits A associates with acting illegally, $\Gamma(\varkappa)$, and inversely with A's associated costs, $\tilde{C}(\hat{\varkappa})$. The conditional arrest probability also increases as P, the actual probability of arrest, increases, but the amount it increases depends on the relationship between objective and subjective probabilities. As we discussed before, an $\alpha(\varkappa)$ greater than 1 implies that subjective probabilities are smaller than objective probabilities, and an $\alpha(\varkappa)$ less than 1 implies that subjective probabilities are greater than objective probabilities. Only if $\alpha(\varkappa)$ equals 1 are subjective and objective probabilities the same.

It is generally agreed that the determination of offense levels and police effectiveness are jointly determined by the interaction of individuals responding to causal and control forces and by the actions of society to control violators. To model police effectiveness, we assume the customary production relationship:[14]

$$P = P(y), \tag{9}$$

where P is the objective probability of arrest and the y are inputs used in detecting and arresting drunken drivers, such as the size of the police force. The inputs represented by y will vary by community since expenditures on y reflect social policy decisions that differ from community to community. Equation (9) can be incorporated in our behavioral relationship, yielding:

$$P(AR|\varkappa,\hat{\varkappa}) = \frac{\Gamma(\varkappa)}{\tilde{C}(\hat{\varkappa})} P(y)^{1-\alpha(\varkappa)}. \tag{8}$$

This relationship provides the basis for our estimation.

III. DATA

In an ideal world we would have an unbiased sample of the population in order to discover how the probability of arrest is related to individuals' characteristics. We do not have that sort of data. Instead, we have data on every arrest for drunken driving in Sweden for the four years 1976-79. For each arrest we know the age and sex of the arrestee and whether there was a previous arrest for drunken driving or other criminal activity. We also know the penalties imposed (fine, jail, license withdrawal, etc.).

We would like to measure $P(AR|\varkappa)$, the probability that a person with a group of characteristics, $\varkappa$, will be arrested. Ideally, we would measure this by $\frac{\#(\varkappa,AR)}{\#(\varkappa)}$, where $\#(\varkappa,AR)$ is

[14] Typical examples of estimates of law enforcement productivity are Darrough and Heineke (1978) and Votey and Phillips (1972).

the number of people arrested with characteristics $\varkappa$, and $\#(\varkappa)$ is the total number of people in the population with characteristics $\varkappa$. Thus, if 100 white, male truckdrivers over age 35 are arrested [$\#(\varkappa,AR)$] and there are 100,000 white, male truckdrivers over 35 in the population [$\#(\varkappa)$], the probability that a white, male truckdriver over 35 will be arrested [$P(AR|\varkappa)$] is .001. Unfortunately, with our data we are forced to estimate the population value, $\#(\varkappa)$, because the figure is not known precisely. We do so by assuming that $\#(\varkappa) = Q(\varkappa) \cdot E$, where $Q(\varkappa)$ is our estimated value of $\#(\varkappa)$ and E is a random error term with lognormal distribution. It then follows that

$$P(AR|\varkappa) = \frac{\#(\varkappa,AR)}{Q(\varkappa)\cdot E}.$$

For convenience we define the measured part of the arrest probability as

$$\Pi(AR|\varkappa) \equiv \frac{\#(\varkappa,AR)}{Q(\varkappa)}.$$

Thus,

$$P(AR|\varkappa) = \frac{\Pi(AR|\varkappa)}{E}. \tag{10}$$

The only personal characteristics, $\varkappa$, available to test our model are age, the number of previous convictions for drunken driving and other offenses, and prior sentences. While we also know the driver's sex, the relatively small number of females arrested for drunken driving and the potential complications that the use of this variable would introduce mean we must confine our study to men only.[15]

We divided the 46,000 male arrestees in our sample into three categories: 25 and younger, 26 to 55, and over 55. These categories are represented by dummy variables; $AG1$, $AG2$, and $AG3$. Our history of drunken driving arrests commences in January 1970 and allows us to characterize those arrested and convicted as having no previous arrests, one previous arrest, or more than one previous arrest. These possibilities are

[15] It is quite clear that the sex variable is important. Of the over 50,000 arrestees in Sweden over the four years of our data, only 4,000 were female. One would expect that a dummy variable for female would be significantly negative, but we were concerned that the differences in drunken driving decisions between sexes was more complicated than the inclusion of a single dummy variable would imply. For instance, if there are two people in an automobile stopped for drunken driving, it is highly likely that the driver, and the one arrested, is a male, even though the passenger is a female and intoxicated as well. It is possible that arresting officers may treat male and female offenders differently.

represented by the dummy variables O if there is one prior conviction and M if there is more than one.

The age variable is intended to capture age-specific tastes as well as attitudes towards the risk associated with drunken driving. The previous conviction variable could enter in various ways. First, one or more previous convictions might indicate a person who enjoyed taking risks, one who was an alcoholic and for this reason often acted irrationally, or one who placed a particularly high value on being able to drive after drinking or saw relatively few additional costs entailed by an arrest. Past arrests might also color an offender's perception of the likely severity of future penalties. Since penalties for drunken driving in Sweden escalate with the number of previous convictions, those with past convictions might be more reluctant to drive while intoxicated than those with no record. Finally, exposure to the criminal justice system may alter a person's view about the costs and probability of an arrest. One might think perceived probabilities should increase with arrest, but both probability and cost judgments may go either way because those who have never been arrested may unrealistically overestimate the likelihood or pain of arrest.

We can adjust our model to take account of the effects of one or more past arrests by expressing the α in Equation (6) as a linear function of previous arrests, i.e.,

$$\alpha(\varkappa) = \alpha_0 + \alpha_1 O + \alpha_2 M. \qquad (11)$$

This implies that the only personal characteristic that affects α is the number of previous arrests. According to this formula the subjective arrest probability of a person with no previous arrests is P^{α_0} ($O = M = 0$). If the person has one previous arrest ($O = 1, M = 0$), the subjective probability is $P^{\alpha_0 + \alpha_1}$, and if a person has been arrested two or more times ($O = 0, M = 1$), it is $P^{\alpha_0 + \alpha_2}$. If α_1 is negative, then people who have been arrested once have a larger subjective probability of arrest than people with no arrest record. If α_1 is positive, then those with one conviction assign a lower probability to being arrested than do those with no record. The subjective probabilities of those with multiple arrests can be similarly compared to those with no arrest. Finally, the relative size of α_1 and α_2 can be used to compare the subjective probabilities of the one-time and multiple arrestee. If α_1 is smaller than α_2, the once arrested have larger subjective probabilities than the multiply arrested, and if α_1 is larger than α_2, those with multiple arrests have higher subjective probabilities than the once arrested.

https://doi.org/10.2307/3053448 Published online by Cambridge University Press

Environmental variables also affect the arrest cost, C. One important factor which may vary by community is the expected punishment in the case of an arrest. Punishments fall into three (not mutually exclusive) major categories: jail, fines, and license withdrawal. For jail costs we use the community-specific probability that a person arrested in a particular month will receive a jail sentence times the average sentence length in days. The expected costs of fines and license withdrawal are similarly defined probabilities conditioned on arrest. We also include these values lagged by one and two months to take account of the possibility that month-to-month fluctuations in these expected sanction costs are sufficiently large to be widely perceived.[16]

Other environmental variables are likely to affect costs of drinking and driving. These include the nature of the transportation system, i.e., distances to be traveled, available modes of transportation, weather conditions, and the like. In addition, local preferences for alcohol might affect the perceived costs of the four ways that drinking and driving can be combined. Preliminary analyses that included proxies for local drinking habits, distances to be driven, vehicle mix, and weather were found to have little impact on our results, so these variables were discarded at an early stage.

The objective probability of arrest is also an environmental variable, since the objective probability varies over the communities of our study. We cannot measure this directly, but we expect that it is a function of the law enforcement resources dedicated to traffic control. Because of data limitations, we represent differences in law enforcement resources by a single proxy variable: hours of police manpower deployed for patrol, as a fraction of total population.[17]

---

[16] One way to adjust for this is to use a moving average (perhaps six months) of the community expected sanctions. A more general way is to use as explanatory variables not only contemporaneous values of the sanctions but also the values of a number of previous months (lagged values). The moving average is a special case of this, if the coefficients on the contemporaneous and lagged variables are all $(1/N)$, where N is the number of months used in computing the average. In practice we found that values lagged more than two months were insignificant.

[17] This is a common proxy in law enforcement effectiveness studies for the simple reasons that other measures of law enforcement inputs are generally unavailable and that, in any event, manpower costs commonly amount to more than 85% of all enforcement resource costs. While in this case one might standardize for community differences in load, say by using driving age population, data on the age distribution of the population were unavailable and, furthermore, there was little reason to believe it varied among the major cities. Other standardizing variables might include traffic density, vehicle numbers, or road mileage patrolled. Such environmental

In summary, the independent variables used in our study are as follows:

$AG1$: 1 if age $\leq$ 25, 0 otherwise
$AG2$: 1 if 26 $\leq$ age $\leq$ 55, 0 otherwise
$AG3$: 1 if 56 $\leq$ age, 0 otherwise
O: 1 if one previous conviction, 0 otherwise
M: 1 if more than one previous conviction, 0 otherwise
$JAIL_i$: Jail cost, lagged i months
$LICWITH_i$: License withdrawal cost, lagged i months
$FINE_i$: Fine cost, lagged i months
$POLICE$: Police patrol hours, per capita

Finally, we account for possible cross-sectional variation unaccounted for by the cost and enforcement variables by using the dummy variables

$STOCK$: 1 if arrest is in Stockholm, otherwise 0
$GOTE$: 1 if arrest is in Götenberg, otherwise 0
MAL: 1 if arrest is in Malmö, otherwise 0
ROC: 1 if arrest not in any of the above three cities, otherwise 0

The dependent variable for this study is the probability that a person with a given set of characteristics is arrested for drunken driving. This is measured by the ratio of two numbers. The numerator, $\#(\varkappa, AR)$, which is the number of people possessing the set of characteristics $\varkappa$ who are arrested in a given month, is observed directly. For every month from 1976 through 1979 we know the number of men in each city convicted of drunken driving, and we can group them by age and prior record. The denominator is not directly observed and must be constructed. We estimate the male population for each city by assuming that the nationwide age distribution given in the *Statistical Yearbook* is uniform across cities. Monthly figures are estimated by interpolating from yearly population figures, which are known for each city, and assuming that year-to-year changes occur smoothly across the months. Unfortunately, there are no available statistics on how many males in each age group have previous arrests. It was necessary to construct these values from our data.

We first computed the average number of men of every age from 15 to 80 arrested for the first, second and more times, by month (Jan., Feb....) for the four years and four regions of our sample. We assumed that these computed averages were the monthly arrests for each age and region since 1970. These

variables were found, early on, not to contribute to the explanatory power of the analysis.

averages were treated as flows into the stock of previously arrested.

Consider the computation for Stockholm drivers age 25 and under. Since our data contain arrest histories that begin in 1970, we start by assuming that on January 1, 1970, no Stockholm driver of age 25 or under had a recorded arrest. By the end of the month some drivers in this group had been arrested and the number is known. Thus, we can specify a stock of arrest records equal to the January average for the Stockholm 25 and under group and a stock of those with no arrest records. By the end of February the stock of those with arrest records can be augmented by the known February average. The stock of men aged 25 can be calculated by adding to the population base (1/12)th of those in the year younger group whose birthdays are assumed to fall in February, and subtracting from the base the number who celebrated birthdays that moved them to the next age group, as well as the number who died (estimated with age-specific mortality rates). We also subtracted from the stock of once arrested the estimated number of individuals arrested a second time. We proceeded in similar fashion to estimate the other age/region combinations for zero, single, and multiple arrests.[18]

With these estimates of crucial terms, we can define from (10) a dependent variable, $\Pi(AR|\varkappa)$, as a ratio of males for each city in each month who are arrested with a specific record (of zero, one, or more than one prior arrest) over the estimated number of all males in that city, in that month, with the same

[18] As an example of how the calculations were made, the formula is presented for the number of 25-year-old males in the various previously arrested categories (zero, one, multiple). Let $POP(25,1,76)$ be the population of males age 25 in January 1976; let $V(1,K,I)$ and $V(2,K,I)$ be the calculated, average number of age K males arrested in Month I for the first time and second time, respectively; let $m(J)$ be the mortality rate for age J, and let Y denote relevant year. The number of multiply arrested males age 25 in January 1976 is estimated for each city

$$M(25,1,76) = \sum_{I=1}^{12} V(2,18,I) + \sum_{I=1}^{12} V(2,19,I) + \ldots + \sum_{I=1}^{12} V(2,24,I)$$

$$= \sum_{J=18}^{24} \sum_{I=1}^{12} (V(2,J,I) - m(J-1)M(J-1,I-1,\Upsilon)).$$

The number of once arrested males is

$$O(24,1,76) = \sum_{J=18}^{24} \sum_{I=1}^{12} (V(1,J,I) - m(J-1)O(J-1,I-1,\Upsilon)) - M(25,1,76).$$

And the number of never arrested is

$$Z(25,1,76) = POP(25,1,76) - O(25,1,76) - M(25,1,76).$$

These values of $M()$, $O()$, and $Z()$ correspond to the $Q(\varkappa)$ variable in Equation (10).

prior record.[19]

IV. ECONOMETRIC SPECIFICATION AND RESULTS

Our examination of the implications of our data starts with Equation (8) combined with (9). Substituting our constructed dependent variable Equation (10) for the true value and taking logarithms[20] indicated by lower case letters, we get

$$\pi(AR|\varkappa) = \gamma(\varkappa) + \beta(\varkappa) + c(y) + (1-\alpha(\varkappa))p(y) + e, \quad (12)$$

where γ is the perceived utility of drinking and driving, β is the perceived cost of an arrest, and c and p are logarithms of actual costs and arrest probabilities as they are determined by the penalty structure and enforcement resources of each community. The perceptual γ and β depend on personal characteristics $\varkappa$. In principle, it is impossible to separate these. However, we have a reasonable idea of the form of the relationship between the characteristics we measure and preferences. For this part of the equation we hypothesize that:

$$\beta(\varkappa) + \gamma(\varkappa) = \delta_0 + \delta_1 AG_1 + \delta_2 AG_2 + \delta_3 O + \delta_4 M. \quad (13)$$

In other words, we assume that a person's judgment about whether the costs of drunken driving outweigh the benefits may be predicted from the person's age and prior record.

As we have noted, $c(y)$ and $p(y)$ are the logarithms of actual costs and probabilities of arrest as determined by the enforcement environment, y. Recall, however, that we are treating these variables in our model as proxies for the expected costs and/or as joint determinants of probabilities of arrest insofar as both expectations and objective values respond to the ecology of enforcement. We hypothesize that the expected costs of an arrest are log linearly related to the expected sanctions where the penalty variables are in their log form. The sanction variables—jail, license withdrawal, and fines—are assumed to discourage drunken driving by making it costly. People, however, may not be well informed of the severity of contemporaneous sanctions, and they may form expectations by what has happened in the recent past. As a proxy for this information process, one and two month lagged

---

19 As in Equation (10), this ratio as constructed measures the true value of the dependent variable with error, i.e.,

$$\Pi(AR|\varkappa) = P(AR|\varkappa) \cdot E.$$

However, since it is the dependent variable, the measurement error causes no problem as long as it is not correlated with the independent variables.

20 We use lower case letters to indicate the logarithm of the respective values. The reader who wishes to examine the development of our model in detail should note that, anticipating the issue of functional form, we have done this throughout.

values of the observed sanction variables are included in the regression along with contemporaneous values. Thus,

$$c(y) = \delta_{50} jail_0 + \delta_{51} jail_1 + \delta_{52} jail_2 + \delta_{60} licwith_0 + \delta_{61} licwith_1 + \delta_{62} licwith_2 + \delta_{70} fine_0 + \delta_{71} fine_1 + \delta_{72} fine_2. \quad (14)$$

The arrest probability, $P(y)$, is treated as an exponential function of the intensity of traffic law enforcement, as measured by police time devoted to traffic law enforcement, holding population constant. Here we hypothesize the typical exponential relationship

$$P(y) = POLICE^{\epsilon}. \quad (15)$$

This form of the objective probability function is a law enforcement production function, the output of which is arrest probabilities. Past estimates agree with the reasonable assumption that ϵ is positive.[21] Assuming ϵ is positive allows us, through the model, to examine whether exposure to the criminal justice system changes a person's perception of arrest probabilities.

A person never arrested for drunken driving (even if he had driven while intoxicated) might conclude, on the basis of his experience, that arrest probabilities are smaller than they actually are. In order to understand how an arrest for drunken driving affects beliefs, recall from (6) that the subjective probability of arrest is

$$\tilde{P}(AR) = P(y)^{\alpha(\varkappa)}.$$

By definition $p(y) = \log P(y)$, and making use of the definitions of (11) and (15), we have for the fourth term on the right-hand side of Equation (12)

$$(1 - \alpha(\varkappa))p(y) = (1 - \alpha_0)\epsilon \cdot police - \alpha_1 \epsilon\, O \cdot police - \alpha_2 \epsilon M \cdot police, \quad (16)$$

the estimates of which will yield coefficients relating subjective probabilities of arrest to prior conviction experience. Since ϵ is positive, the sign on the coefficient, along with the relative magnitudes, of police, $O \cdot police$, and $M \cdot police$ will reveal the relation between subjective and actual arrest probabilities.

With our formulation, the subjective probability of arrest is

$$\tilde{P}(AR) = P^{\alpha_0 + \alpha_1 O + \alpha_2 M}. \quad (17)$$

Therefore, for the never arrested ($O = M = 0$), the subjective probability is

$$\tilde{P}(AR) = P^{\alpha_0}. \quad (18)$$

[21] Votey and Phillips (1972) find estimates of its value to be in the neighborhood of 0.75, using the exponential form typical of much estimation of productivity relationships.

https://doi.org/10.2307/3053448 Published online by Cambridge University Press

And since P is smaller than 1, raising it to a power less than 1 increases its value. If, for instance, the actual probability of arrest, P, $= 1/4$ and $\alpha_0 = 1/2$, then the subjective probability of arrest, $\tilde{P}$, $= 1/2$. If, however, $\alpha_0 = 2$ (when $\alpha > 0$, the coefficient on *police* is negative), then $\tilde{P} = 1/16$. To carry this example further, suppose the coefficient on $O \cdot$ *police* is positive, which implies that α_1 is negative. In this case, $\alpha_0 + \alpha_1 O$ is smaller than α_0, meaning that the once arrested assign a higher probability to being arrested than do the never arrested.

In summary, if $\alpha_0 = 1$ or if $\alpha_0 + \alpha_1 O$ or $\alpha_0 + \alpha_2 M$ sum to 1, then the subjective probabilities equal the actual probability. Values greater than 1 imply that individuals underestimate the actual probability, and values smaller than 1 imply overestimation. Furthermore, the larger the values, the smaller the probability assessment. Therefore, from the regression results it will be possible to gauge how subjective estimates of the probability of arrest vary with arrest records. This will be reflected in the estimated values for the α_i.

This method of analysis may also yield information about whether apparent deterrent effects reflect specific or general deterrence. Any implications derived from the estimation of α will be individual-specific (i.e., based on personal experience rather than on perceived general risks) since what we are measuring is a change in behavior (learning) associated with the unpleasant experience of arrest and conviction. However, perceived costs are also influenced by the general levels of enforcement and sanctions observed in the community through $c(y)$ from (14). Effects captured here are potentially independent of personal experience, depending upon the way the relation is estimated. We shall, however, not be able to separate specific from general deterrence effects because of the nature of our data.

V. SUMMARY OF RESULTS

Equation (12), incorporating specifications (13), (14), and (16), was estimated with ordinary least squares. The results are given in Table 1. They suggest the following relationships. Men in the middle age group are significantly more likely to drink and drive than are either younger or older men. The most dangerous age for drunken driving is 26 to 55. The probability that men in this age group will drink and drive is 1.2 percent higher than it is for men in the younger group, the omitted category, and 4.7 percent higher than it is for men in the older group. Other things being equal, the probability that

https://doi.org/10.2307/3053448 Published online by Cambridge University Press

men who are older than 55 will drink and drive is 4.6 percent smaller than the comparable probability for those younger than 26. Men who are arrested for drunken driving are less likely to drink and drive in the future. One arrest for drunken driving lowers the probability of a future drunken driving incident by 6.8 percent and a person with more than one arrest is 16.4 percent less likely to drink and drive in the future.

Table 1. Regression Results

Variable	Coefficient	Standard Error
Constant	8.476	8.446
$AG2$	1.237*	0.266
$AG3$	−4.671*	0.266
O	−6.754**	2.859
M	−16.436*	2.859
$police$	1.662**	0.740
$O \cdot police$	−1.683**	0.644
$M \cdot police$	−3.762*	0.644
$licwith_0$	−0.591	0.442
$licwith_1$	−0.283	0.463
$licwith_2$	−0.881***	0.439
$fine_0$	−0.140	0.765
$fine_1$	0.019	0.754
$fine_2$	0.085	0.724
$jail_0$	−0.457	0.474
$jail_1$	0.182	0.522
$jail_2$	−0.588	0.500
$STOCK$	−1.371***	0.700
MAL	−2.628*	0.375
$GOTE$	−2.715*	0.647

R^2: 0.353

$D.W.$: 1.876

*Significant at 1% level.
**Significant at 3% level.
***Significant at 5% level.

We cannot fully explain these results but we have some conjectures. The first is that people really don't know how bad, or how costly, an exposure to the criminal justice system is until they have been arrested. The problem with this explanation is that it suggests that one exposure is not enough to learn fully how unpleasant exposure is. Second, Swedish law penalizes second offenders more than first offenders and multiple offenders more than second offenders. Thus, we may have spotted a pure deterrent impact, reflecting differences in the perceived severity of the punishment likely to accompany the next offense. Third, the result might be due to our overestimating the stock of one-time and multiple offenders. In constructing these values, we assumed that the mortality rate in these groups was the same as in the rest of the population. However, these rates may be higher for those who drink and drive than for those who don't. We experimented with different assumed mortality rates, and reasonable changes made little difference in the statistical results. Finally, the result may be due to the effects of prior sanctions. If during the period spanned by our data a significant number of those previously arrested were not driving (because they either were in jail or had had their licenses taken away), our estimate of the stock of these people would be too large. This would cause the coefficients O and M to be biased downward. A similar effect would occur if prior sanctions, such as placement in an antiabuse program, reduced drinking apart from driving behavior.

The implications of the estimated coefficients on $O \cdot$ *police* and $M \cdot$ *police* are particularly interesting because of what they suggest about subjective probabilities. As we mentioned before, a value of $\alpha > 1$ implies that people overestimate arrest probabilities, and a value of $\alpha > 1$ implies that they underestimate these probabilities. Although we cannot estimate ϵ directly, because with this specification the coefficient represents $\epsilon(1 - \alpha)$, for reasons we have mentioned, it is reasonable to expect that ϵ is positive. Our results indicate that first offenders at the time of their arrest have substantially higher estimates of the probability of being arrested than individuals with one prior arrest and that people with one arrest have higher estimates than those with two or more. Put another way, multiple offenders, if acting rationally, apparently place a much lower probability on being caught than the general population.

These results may be read as confirming the view that many repeat offenders are not rational in their evaluation of the risk of punishment or, if they can appreciate the risk, cannot rationally perceive its behavioral implications. These kinds of irrationality might be thought especially likely in a group disproportionately populated by alcoholics. On the other hand, the results are not necessarily inconsistent with rational behavior. Multiple offenders may have a special taste for risk or they may find particular utility in driving after drinking. More interestingly, they may, particularly if the group is dominated by alcoholics, frequently engage in driving after drinking. If they usually escape arrest while doing so, they may more accurately assess the (low) risk of apprehension for driving after drinking than those who seldom engage in this kind of deviance and so have never been arrested or have been arrested just once. Another striking result is that the costs that one would expect to be associated with the perceived probability of jail and fines do not have a significant impact on the probability of arrest. The current and lagged coefficients on these penalty variables are all statistically insignificant. Costs associated with the chance that a driver's license will be withdrawn do appear to have a significant lagged effect. That the effect is lagged can be explained by the delay in dealing with license revocations, which are handled as separate decisions from convictions. Because of this delay, it is not surprising if people's perceptions of the rate at which licenses are being withdrawn reflects decisions made some months before.[22]

VI. CONCLUSIONS

Rational choice models assume that as the cost borne by criminals increases, law violations decrease. Our results suggest that such models may be applied to the crime of driving after drinking, a crime that in the minds of some is peculiarly rooted in irrational behavior. Our results are also consistent with models of learning in which repeated exposure or treatment enhances learning and/or tends to modify behavior.

A rather striking facet of these results is their consistency with the belief that, in attempting to control driving after drinking, we are dealing with two disparate populations, one that is capable of learning and one that persists in ignoring the law and the extent of penalties. We see this in the juxtaposition of the fact that multiple offenders are less likely to be arrested than those with zero or one prior arrests with the fact that multiple offenders who are arrested apparently perceive a low probability of arrest or see the benefits of

[22] These results might be interpreted as general deterrence effects, had we obtained our results in such a way as to rule out the effects of experience in coloring the information associated with perceptions of general probabilities or costs. To do so would have required separate estimates for each class of arrestee, i.e., those without priors, and with one or more. Our estimates are average coefficients over all arrestees for the cost variables.

We didn't pursue this option because of the insignificant result on jail and fine costs and limitations on resources for estimation. However, it would be premature to suggest that general deterrence doesn't work, except for driver's license withdrawal. The reason for this is the high degree of uniformity in sentencing that prevails in Sweden. Swedish researchers had cautioned us about the likelihood that our series on penalties would not show sufficient variance for significant estimates. Our results on the cost variables suggest this to be true, except for license withdrawal.

driving while drunk at levels that for most people would be irrationally high. This interpretation of our results is consistent with the observation that many of those in the population of violators are alcoholics or intransigents who will not or cannot conform to the laws under any reasonable circumstances, while others respond to the laws in predictable ways, making the imposition and enforcement of drunken driving laws an important element in social control.

Our analysis is, as we stated at the outset, limited by the nature of the available data. We have had to estimate key variables with the potential error that entails, and while our model calls ideally for a rich vector of personal traits with which to pinpoint subjective costs and probabilities, we have had to work with only a few. Nevertheless, the results our approach yields are plausible. More importantly, we think we have demonstrated a viable way of overcoming some of the difficulties that have plagued prior deterrence research that attempts to test an individual level theory with models of aggregate behavior.

REFERENCES

BECCARIA, Cesare (1963) *On Crimes and Punishment*. Indianapolis: Bobbs-Merrill.

BENTHAM, Jeremy (1948) *The Principles of Morals and Legislation*. New York: Hafner.

BLUMSTEIN, Alfred J., Jacqueline COHEN and Daniel NAGIN (1978) *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*. Washington, DC: National Academy of Sciences.

DARROUGH, M.N. and J.M. HEINEKE (1978) "The Multi-Output Translog Production Cost Function: The Case of Law Enforcement Agencies," in J.M. Heineke (ed.), *Economic Models of Criminal Behavior*. Amsterdam: North Holland.

HEINEKE, J.M. (ed.) (1978) *Economic Models of Criminal Behavior*. Amsterdam: North Holland.

INTERNAL REVENUE SERVICE (1980) "A Procedure for Estimating Taxpayer Response to Changes in IRS Audit Coverage," submitted to IRS by Westat Inc.

ISRAELSTAM, S. and S. LAMBERT (eds.) (1974) *Alcohol, Drugs and Traffic Safety: Proceedings of the 6th International Conference*. Toronto: ARF Books.

JONES, Ralph K. and Kent B. JOSCELYN (1978) *Alcohol and Highway Safety 1978: A Review of the State of Knowledge*. Ann Arbor, MI: Highway Safety Research Institute.

NORSTRÖM, Thor (1981) *Studies in the Causation and Prevention of Traffic Crime*. Stockholm: Almqvist & Wiksell International.

PHILLIPS, Llad (1978) "Factor Demands in the Provision of Public Safety," in J.M. Heineke (ed.), *Economic Models of Criminal Behavior*. Amsterdam: North Holland.

PHILLIPS, Llad, Subhash RAY and Harold L. VOTEY (1984) "Forecasting Highway Casualties: The British Road Safety Act and a Sense of Déjà Vu," 12 *Journal of Criminal Justice* 101.

ROSS, H. Laurence (1981) *Deterrence of the Drinking Driver: An International Survey*. Washington, DC: National Technical Information Service.

SHAPIRO, Perry (1980) "A Model of Tax Evasion." Mimeo.

VOTEY, Harold L. (1984) "The Deterioration of Deterrence Effects of Driving Legislation: Have We Been Giving Wrong Signals to Policymakers?" 12 *Journal of Criminal Justice* 115.

VOTEY, Harold L. and Llad PHILLIPS (1972) "Police Effectiveness and the Production Function for Law Enforcement," 1 *Journal of Legal Studies* 423.