DOES REPORTING DETER BURGLARS?

An Empirical Analysis of Risk and Return in Crime

I. Goldberg and F.C. Nold

October 1978
Technical Report CERDCR-4-78

Prepared under Grant No. 77-NI-99-0071 from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice.

CENTER FOR ECONOMETRIC STUDIES OF THE JUSTICE SYSTEM

Hoover Institution, Stanford University

59663

I. Introduction

Recent studies have analyzed the demand for private protection against crime[1] and discussed the interaction between private security and public safety. The argument that cooperation by the public with the police is required to successfully combat crime has also been frequently raised in popular debates. However, while the deterrent effect of public law-enforcement has been studied extensively in the literature, systematic empirical studies of the effect of private enforcement on crime are not abundant.[2]

Private protection can take many forms; self-protection devices and other protective behaviors are probably the most frequently mentioned examples. Our emphasis is on another form of private behavior that nevertheless may substantially affect the success of public law enforcement: the reporting of crimes to the police by victims. Police effort to apprehend and convict criminals crucially depends on the reporting of crimes, and the lack of cooperation by victims is often blamed for police failure to apprehend and convict criminals. Because reporting increases the chances of apprehension and conviction, it should have a deterrent effect on potential offenders. To study this effect we chose to concentrate on burglary in residences. Our basic premise is that potential residential victims who are perceived by burglars to have a high tendency to report are less attractive as burglary targets. This perceived reporting probability of a given household thus become a victim-specific

[1]See, for example, Clotfelter (1977)

[2]An important exception is Clotfelter (1978); however, due to unavailability of data, he could not test empirically the effects of private protection on crime. Komeasar (1973) analyzes some determinants of victimization, using a National Opinion Research Center study, but abstracts from the deterrence question. The demand for private protection is also analyzed by Bartel (1975).

deterrent variable.

Apart from our concern with private enforcement, this study is an empirical re-investigation of the deterrence question by using individual observations on victims and potential victims. We are among the first to use National Crime Panel (NCP) victimization surveys. These include information about the victimization experience of three hundred thousand households, fourteen per cent of which were burglarized in the year preceding the interview. The surveys reveal that only about half of all burglaries are reported to the police.

These data have four advantages vis-a-vis the aggregate crime statistics used in former studies. First, they allow us to construct a victim-specific deterrent variable, rather than to assume that all victims in a state or city impose on the offender the same amount of risk. Second, information on losses from burglary enables us to use a direct measure of the illegal returns. Third, in contrast to police-recorded crime statistics, the NCP surveys include information on all crimes, either reported or not. Fourth, the use of individual data makes it easier to avoid the simultaneity problem encountered in studies based on aggregate data. The present study offers, in our view, a more direct and reliable test of the deterrence hypothesis.

II. Analytical Framework

The model of time-allocation between legal and illegal activities developed in the literature (Becker, 1968 and Ehrlich, 1973) is applied in this section to the choice between burglary and alternative activities. Assume that the potential burglar[3] faces a set of households and his problem is to choose a subset of them as burglary targets on the basis of his expected returns. We first describe in this section how the potential burglar forms his expectations about these returns. At the end, we discuss his time-allocation decision and the parameters that determine the probability of victimization.

The net expected return from burglarizing the ith household is:

(1) $B_i = B_i(L_i, p_i^R, p_i^C, F_i)$

where L_i is the expected gross return ("loot") from the burglary of the ith household

p_i^R is the probability that the ith household will report the burglary to the police[4]

p_i^C is the probability that the burglar will be apprehended and punished for the burglary of the ith household

F_i is the level of punishment for the burglary of the ith household.

Obviously, the expected gain rises with the magnitude of the "loot" L_i. The other three variables are components of the expected cost which is imposed on the offender as a consequence of burglarizing the ith household. We may realistically assume that apprehension and conviction of burglars depend upon the reporting of the crime to the police. These deterrence variables are therefore interdependent; they are entered separately into the B_i function (instead of entering $p_i^R \cdot p_i^C \cdot F_i$ as one variable) since in general

[3] All potential burglars are assumed to be identical.

[4] Reporting by persons other than the victim is disregarded since the surveys show that only a trivial percentage of burglaries are reported by others.

their respective partial effects are not necessarily of the same magnitude although all three effects are assumed to be negative.

In order to predict the value of B_i for each household in his choice set, the offender must possess some knowledge about the values of the variables in the B_i function. The first two variables, L_i and p_i^R, are the more difficult to predict since they depend on the individual behavior and characteristics of each household. We assume the offender predicts these a priori unknown values on the basis of observable household characteristics. For L_i we write:

(2) $L_i = L_i(W_i, T_i)$

where W_i is the victim's wealth and T_i is the cost of burglarizing one dollar value from household i. The parameters that determine this cost depend on choices made by the household, choices that are only in part related to possible victimization. For example, if W_i is held in an easily accessible and transferrable form, T_i is lower. Also, the type of the housing structure (e.g., single-unit house versus apartment), impediments to entry (locks and bars), alarms and watchdogs determine the level of T_i. Moreover, it is plausible that the presence of persons at home while a burglary is being attempted increases T_i.[5] The probability that at least one household member is home at any point of time is a function of family structure, behavior patterns and--in some cases--even a deliberate choice to stay home in order to protect one's

[5] The victimization surveys show that burglaries attempted when somebody is present at home result in a lower loss.

property.[6] Since this presence probability can be deduced from observable traits (see Section III) and other determinants of T_i are also partly observable, the burglar can predict T_i.

Regarding the prediction of the household's reporting probability, let us postulate:

(3) $p_i^R = p_i^R(L_i, C_i)$

where C_i is the cost of reporting. Reporting behavior is thus postulated to be determined by the gain that the household expects due to reporting and the cost of doing so. The offender, in turn, bases his prediction of p_i^R on the observed differential gains and costs across households. The household gains from reporting by recovery of losses, and since the main recovery channels--insurance and tax-deduction--require that the burglary be reported to the police, L_i should have a positive effect of p_i^R.[7] We hypothesize that the main component of the cost of reporting is the opportunity time cost; the latter may be non-negligible since reporting has to be done in person and in some cases may involve police visits, trips to police headquarters and days in court.

The household is certainly aware of the fact that its reporting behavior affects its attractiveness as a burglary target; should this be a factor in

[6] Such behavior represents an additional cost of crime (see Clotfelter, 1977).

[7] To the extent that physical recovery by police is important we should include p^C in the p_i^R function. Since this recovery rate is only five per cent, we disregard this effect. (Empirically, p^C was found to have no effect on reporting. See footnote 15.)

its reporting decision? We think it realistic to assume that past reporting behavior of an individual household is unknown to the offender. Although he is aware of the differences in reporting behavior among different household types, we rule out the possibility that the burglar traces back the reporting history of any specific household. He makes his prediction of p_i^R solely on the basis of observable household traits because the household has no credible way to signal its reporting intentions to potential burglars;[8] therefore the potential victim can rationally ignore the effect of p_i^R on B_i in making his reporting decision. However, our approach to the reporting decision is still consistent with a community effort to deter crime by increasing its reporting. Although it is unrealistic that individual past reporting behavior is observable, it is reasonable that differences in reporting behavior among groups of individuals (neighborhoods, cities, ethnic groups, etc.) are observable by burglars. Therefore, if all households in a group get together and decide to report all burglaries, their behavior can provide a signal to offenders. (Obviously, though, such group decision encounters the well-known "free rider" problem.) Our empirical analysis of individual observations is mostly concerned with the victim-specific deterrent effect, but the analysis of aggregate deterrence bears some relevance to community behavior.

The offender's learning process about the public enforcement variables

[8] Announcement of a household's reporting intentions, such as signs saying "We report all crimes," is not credible because people can post such signs without actually altering their reporting behavior. The argument that other behavior, such as signs saying "All valuables marked," indirectly indicates reporting intentions is tenuous and difficult to justify on theoretical grounds.

p_i^C and F_i is assumed to be rather simple. Even though in principle these parameters are also target-specific,[9] he is assumed to estimate their values by learning the <u>average</u> level of apprehension rates and penalties in the relevant subset of population (e.g., state, city). In contrast to private behavior, knowledge about behavior of the police or courts is easier to obtain, and therefore this assymteric assumption about knowledge of private and public activities is sensible. Moreover, the data available do not allow us any study of target-specific public protection.

The time-allocation model implies that the potential offender equates his marginal value of time--adjusted for risk--in the legal and illegal activities. To apply this marginal condition to our case, let us assume that a constant amount of time is required to burglarize each household in the choice set and this amount is equal across households; for illustration, let it be one hour per household.[10] We denote the marginal (hourly) value of the potential offender's time in an alternative (second best) legal or illegal activity as w. The individual allocates his time to burglary as long as his hourly return in burglary, B_i, exceeds w. He will thus burglarize all households whose B_i exceeds w; i.e., for the ith household:

$$(4) \quad V_i = \begin{cases} 1 & \text{Victimized, if } B_i > w \\ 0 & \text{Not victimized, otherwise.} \end{cases}$$

[9] For example, burglarizing the house of an influential local politician may involve higher p^C and F than burglarizing other households.

[10] This assumption implies that the offender does not adjust the amount of time he spends on a target in response to B_i.

Empirical implementation of this conceptual framework is accomplished by introducing errors in the predictions made by potential burglars. Since both L_i and p_i^R are predicted on the basis of observables, their estimated values are composed of their true values plus error terms. (For simplicity, we assume that F_i and p_i^C are known with certainty.) Obviously, then, the predicted value of B_i is also a random quantity. Therefore, the victimization probability of a household is the probability that the predicted value of B_i (i.e., its true value plus an error term) exceeds w; we denote this probability by $V_i^* = p(V_i = 1) = p(B_i > w)$. Consisently with our previous assumptions, the implications of (4) are:

$$\frac{\partial V_i^*}{\partial L_i} > 0 \qquad \frac{\partial V_i^*}{\partial p_i^R} < 0 \qquad \frac{\partial V_i^*}{\partial p_i^C} < 0$$

$$\frac{\partial V_i^*}{\partial F_i} < 0 \qquad \frac{\partial V_i^*}{\partial w} < 0$$

III. <u>Empirical Analysis</u>

The primary purpose of our empirical work is to estimate a victimization equation designed to test the implications of Section II. Since this equation is to be estimated from a sample of households, the majority of which were not burglarized, our first step is to create predicted values for p_i^R and L_i. This is done by estimating auxiliary loss and reporting regressions for the subsample of burglarized households and using their coefficients to estimate predicted values of p_i^R and L_i for each household in the whole sample of households, either burglarized or not.

Data Base

The main data source is the NCP victimization surveys, conducted by the U.S. Bureau of the Census, which included eight cities in 1972, five cities in 1973 and thirteen cities in 1974. In 1975, the thirteen cities, which were covered in 1972-73, were surveyed again. Each household included in the survey was questioned about its victimization history in the year preceding the interview. Information was obtained about each incident reported, and details about household and individual characteristics were collected. We have constructed a sample of over 300,000 households identified by city. For each of the 14% of the households that were burglarized during the year preceding the interview an incident file, which described the burglary, was appended.

The Loss Equation

Following (2), the explanatory variables in this auxiliary regression are proxies for W_i and T_i. Household income is used as a measure of wealth (precise definitions of variables are given in the appendix). Information on some of the parameters which determine the cost of burglary T_i is available in our data set. We included indicators of the type of housing structure because single-unit houses and irregular structures (trailers, etc.) are more easily accessible targets.[11] Length of residence in the neigh-

[11]The variable RENT is used as a proxy for multi-unit houses. Note that house ownership might also be a proxy for wealth.

borhood increases one's knowledge of local crime conditions and thus affects his protective behavior; RECENT, which measures length of residence, was introduced to capture this effect. Several household characteristics were used as predictors of the probability that at least one person is present at home at any point of time. This probability is higher for households with young children (CHILDREN < 12) or two or more adults; life-cycle patterns of labor force participation and human capital investment suggest a higher probability for older persons, especially the retired, and a lower probability for young people, mostly students. Race was introduced as a control variable.

Table 1 contains a summary of the estimated multinomial LOGIT model[12] for six loss categories.[13] Inspection of the results reveals that, although household income is indeed the main explanatory variable of loss, the other variables generally have reasonable and significant effects, as well. We do not elaborate more on this equation because its primary purpose is the prediction of L_i for the victimization equation.

The Reporting Equation

Following (3), we include in this regression the value of items stolen as a proxy for L_i. In addition, since insurance and tax deductions may provide compensation due to damages incurred, we include the value of the property damage. Obviously, the potential for compensation depends on whether or not the household was insured against theft. House ownership (as opposed to residence in rented apartments) may serve as a signal to the burglar that

[12]Since 40% of the burglaries result in a zero loss, the loss equation cannot be estimated by the OLS method. For technical reasons, we used a multinomial LOGIT model instead of, for example, a TOBIT model.

[13]More detailed results are presented in the appendix.

Table 1

Effect of Explanatory Variables on Loss Structure

	LOSS 0	LOSS 1 - 25	LOSS 25 - 75	LOSS 75 - 250	LOSS 250 - 500	LOSS > 500
INC = 0 - 3000	+*	++	+	++	--	--
INC = 3000 - 6000	-	++	-	+	--	--
INC = 10000 - 25000	--	++	++	++	--	++
INC > 25000	--	+	++	++	-	++
RACE = BLACK	--	--	++	++	++	++
AGE < 25	++	--	++	++	++	++
AGE > 55	++	--	--	+	--	--
CHILDREN < 12	-	++	++	--	--	--
2.ADULTS	--	++	++	--	--	--
RENT	++	--	--	+	--	--
IRREG.HOUSING	--	++	-	-	+	--
RECENT	--	-	++	++	++	++

*The sign indicates the direction of change in the probability. Two signs indicate that the change is statistically significant at the 10% level. The actual magnitudes of these changes appear in Table 2 of the appendix. Explanation of the display procedure is provided in Part B of the appendix.

the victim is insured because many mortgage agreements require property insurance, which usually includes theft coverage; RENT is therefore included in the regression variables in addition to WITH.INS. The available data do not provide us with satisfactory measures of the cost of time of household members. In the absence of wage rates for individual members, we use household income as a measure of time-cost which should have a negative effect on reporting, given loss. Human capital theory suggests that, given the other determinants of wages, a person's wage rate first rises and finally declines with age. Therefore, in addition to income, which partially captures the other wage determinants, we include the age variable. It is introduced in a non-linear form (AGE < 25, AGE > 55) because of its hypothesized non-linear effect on the wage rate.[14] Because burglaries--mainly attempted ones--that occur when a household member is present at home are generally reported, probably because of the threat of violence, the variable PRESENT was introduced to control for this effect. We also introduced an interaction-term between low loss and damages because when the loss sustained is low, property damage provides an indication that a burglary, rather than inadvertent misplacement of some items, has occurred.

Inspection of the results in Table 2 reveals that loss, property damage and insurance have the expected positive effects. Home ownership

[14] It was hypothesized that schooling should have a positive effect on reporting because more educated individuals may be more efficient in reporting and therefore face a lower cost. No significant effect was found.

TABLE 2

Estimated Reporting Model
Maximum Likelihood LOGIT*

Explanatory Variable	Coefficient
INCOME=0-3K	-0.033 (-0.597)†
INCOME=3-6K	-0.103 (-1.719)
INCOME=10-25K	-0.010 (-2.076)
INCOME>25K	-0.266 (-2.943)
LOW.LOSS, DAMAGE	0.452 (4.503)
PROPDAM=1-50	0.382 (7.761)
PROPDAM>50	1.277 (12.353)
LOSS=1-25	0.132 (2.618)
LOSS-26-75	1.316 (20.229)
LOSS=76-250	1.880 (23.204)
LOSS=251-500	2.133 (22.689)
LOSS>500	2.786 (32.831)

*These estimates are for 14 of the cities surveyed in 1973 and 1974.

†The values given in parentheses are the test statistics of the coefficients.

TABLE 2

continued

Explanatory Variable	Coefficient
WITH.INS	0.982 (16.443)
AGE<25	-0.148 (-2.771)
AGE>54	0.125 (2.615)
SOMEONE.PRESENT	0.816 (15.494)
RENT	-0.062 (-1.450)
CONSTANT	-1.008
-2*LN(L)	17783.2

has a positive but hardly significant effect.[15] As for the proxies of the cost of time, the coefficients for the high income classes are negative as expected (the mid-income class INCOME=7-10K was omitted), but the income effect is non-linear because the low-income classes also have negative coefficients, although hardly significant. As expected, AGE > 54 has a positive coefficient but, contrary to our expectation, AGE < 25 has a negative coefficient. The variable PRESENT has a strong positive effect.[16]

The Victimization Equation

This equation was estimated in the whole sample of interviewed households. The dependent variable is a dichotomous victimization-indicator, VICT (precise definitions are given in the appendix). The household-specific explanatory variables are the predicted values from the loss regression (LOSS) and the reporting regression (PREP).[17] The probability of apprehension is approximated by a city-wide variable, CLR; data on penalties are not available for the cities and time-periods of our sample. Although it was assumed in Section

[15] As suggested on page 5, footnote 7, CLR was also introduced in some preliminary work but no significant effect was found.

[16] It was felt that the effect of losses on reporting might be different in different income groups because higher income households might have a different marginal evaluation of the loss. The same regression (without income) was estimated in each income class and the appropriate likelihood ratio test was performed by determining Pr(CHISQUARE(70 - 18) > 17783.2 - 17679.92 = 103.28) where 17679.72 is the sum of the likelihood functions estimated by splitting into income classes and 17783.2 is the value of the likelihood function for the pooled model. This probability is less than .001 and so we conclude that splitting into income classes is a statistically superior procedure. The income classes version shown in the appendix was therefore used to predict the values of p_i^R that were used in the victimization equation.

[17] The reporting regression includes two variables--PROPDAM, SOMEONE. PRESENT--which cannot be known for unburglarized households. (Also, only burglarized households were asked about insurance.) To construct the predicted p_i^R for each household, PREP, we assigned to these unknown variables values which were characteristic of similar burglarized households. Detailed descriptions of the procedure can be found in an appendix available upon request.

[This attachment is included for the referee's use.]

II that all offenders are alike, differential demand conditions and positive migration costs might cause w to vary across locations. To the extent that this is important, we introduced a measure of legal opportunities on a city-wide basis.[18]

Table 3 contains maximum likelihood estimates of a LOGIT model of the probability that a household will be victimized. All the estimated coefficients have the hypothesized signs and all except CLR are statistically significant. The most important result is the negative and significant effect of the reporting probability on the victimization probability. Note also the large elasticity of PREP--roughly 2. We claim that this finding provides empirical support--on a microeconomic level--for the deterrence hypothesis: if a given household is perceived to have a higher potential reporting probability, its victimization probability is lower because potential burglars are deterred by its higher p_i^R.

In Table 4 we use our estimated model to predict the victimization probability of the average household for various potential values of PREP. To interpret these results recall that the gains from reporting and its costs depend on household traits and therefore are different across households. Our results show that households which are characterized by low p_i^R have a higher victimization probability.

Since our hypotheses suggest that the expected value of the "loot" has two opposite effects on the victimization probability, it is interesting to calculate the net effect of expected loot on victimization:

---

[18] We were unable to find satisfactory measures which are compatible with our sample. The variable shown in Table 3, PCPOOR, is hardly satisfactory. A similar measure is used by Ehrlich, 1973.

TABLE 3

Estimated Victimization Model
and Elasticities

(Maximum Likelihood LOGIT)

Explanatory Variable	Estimated Coefficient	Estimated Standard Error	Test Statistic	Elasticity
PREP	-4.701	+0.673	-6.986	-2.165
LOSS	+.0074	+.0012	+6.065	+0.762
PCPOOR	+3.097	+1.235	+2.507	+0.598
CLR	-0.314	+0.334	-0.940	-0.056
CONSTANT	-0.799	-----	-----	-----

The value of minus twice the log of the likelihood function is 8169.07.

The number of observations is 10008. The mean of the dependent variable, VICT, is 0.1414, and the means of the explanatory variables are: PREP 0.5363, LOSS 119.9025, PCPOOR 0.2248 and CLR 0.2086.

TABLE 4

Reporting and Victimization*

	Potential Minimum	Minimum Imputed Value**	Mean of the Sample	Maximum Imputed Value	Potential Maximum
Probability of reporting (PREP) set at:	0.0	.361	.537	.714	1.0
Victimization probability:	.637	.318	.141	.113	.043

*These are based on the coefficients presented in Table 3. All explanatory variables except PREP are assigned their mean values.

**This refers to the extreme values encountered in the sample of 19,716 observations.

$$\frac{dV_i^*}{dL_i} = \frac{\partial V_i^*}{\partial L_i} + \frac{\partial V_i^*}{\partial p_i^R}\frac{\partial p_i^R}{\partial L_i}$$

$$\frac{dV_i^*}{L_i} = .89 \times 10^{-3} - (.57)\ .69 \times 10^{-3} = .50 \times 10^{-3}$$ [19]

It is interesting to note that the net effect of the expected "loot" on victimization probability is still positive[20] but it is almost half of the gross effect. The attractiveness of high loot targets is substantially reduced by their higher perceived tendency to report.[21]

---

[19] Since the logistic function is of the form $V_i^* = e^{BX_i}/(1 + e^{BX_i})$ where X_i is a vector of the explanatory variables (see Table 3),

$$\frac{\partial V_i^*}{\partial L_i} = .0074\ V_i^*(1 - V_i^*) = .89 \times 10^{-3}$$

and

$$\frac{\partial V_i^*}{\partial p_i^R} = -4.701\ V_i^*(1 - V_i^*) = .57.$$

The coefficient $\frac{\partial p_i^R}{\partial L_i}$ was obtained from a re-estimation of the equation in Table 2, with L_i introduced continuously:

$$\frac{\partial p_i^R}{\partial L_i} = .0028\ p_i^R\ (1 - p_i^R) = .69 \times 10^{-3}.$$

[20] It can be shown that the net effect is still significantly different from zero.

[21] One could argue that other forms of private enforcement such as locks, bars or alarms are positively correlated with reporting. If so, it is possible that the estimated effect of PREP on victimization captures the combined effect of all forms of private enforcement, i.e., it has to be interpreted as a deterrent effect in a broader sense. We cannot test this hypothesis since data on protective devices are unavailable.

Aggregate Deterrence

As an alternative to the specification of the equation in Table 3, we subsequently estimated the following equation:

$$p(V_i = 1) = p(\alpha_0 + \alpha_1 \text{ RELPREP} + \alpha_2 \text{ RELLOSS} + \alpha_3 \text{ PCAPTURE} + \alpha_4 \text{ PCPOOR} > 0)$$

where: RELPREP = PREP - PBARCITY
RELLOSS = LOSS - LBARCITY
PCAPTURE = CLR * PBARCITY

and PBARCITY and LBARCITY are the city-wide averages of PREP and LOSS, respectively. This specification attempts to decompose the deterrent effect of reporting into two components: α_1 should capture the effect of the household-specific p_i^R relative to other burglary targets: this intra-burglary deterrent effect represents a shift by burglars from target i to other households. The coefficient α_3 is supposed to capture the deterrent effect which occurs through the shift of offenders from burglary to other activities. Inspection of Table 5 reveals that the aggregate variable PCAPTURE has a negative and significant effect, but its elasticity is low.[22] Since the effects of CLR and PBARCITY

However, theoretical considerations suggest that, for given levels of LOSS, there should be a negative correlation between reporting and use of protective devices because the former is time-intensive while the latter is goods-intensive. Therefore, high income individuals should, given LOSS, report less and use more protective devices. If such negative correlation exists, the omission of protective devices biases the coefficient of PREP toward zero; i.e., the effect would have been stronger if devices had been included in the equation.

[22] Alternatively to RELPREP, we used the ratio form PREP/PBARCITY and the standardized form (PREP - PBARCITY)/SDPREP where SDPREP is the standard deviation of PREP over households in a city. The same was applied to LOSS. The qualitative results were unchanged.

TABLE 5

Estimated Victimization Model
and Elasticities

Explanatory Variable	Estimated Coefficient	Estimated Standard Error	Test Statistic*	Elasticity**
RELPREP	-6.631	+Ø.497	-13.343	-3.Ø6
RELLOSS	+Ø.Ø13	+Ø.ØØ1	+14.Ø58	+1.31
PCAPTURE	-1.223	+Ø.445	-2.136	-Ø.12
PCPOOR	+1.955	+Ø.915	+2.136	+Ø.38
CONSTANT	-2.153	---	---	---

*The estimated coefficient divided by its estimated standard error is presented in this column. These values can be compared to the Standard Normal Density to determine significance levels.
**The means calculated from the random sample of 19716 observations used in estimation of the model are; percent victimized, .14Ø, PREP, .537, LOSS, 119.67 , PCAPTURE, .112, PCPOOR, .225.

-21-

TABLE 6

Estimated Victimization Model
and Elasticities,
An Alternative Specification

Explanatory Variable	Estimated Coefficient*	Estimated Standard Error	Test Statistic**	Elasticity***
RELPREP	-6.378	+0.701	-9.097	-2.94
RELLOSS	+0.013	+0.001	+10.556	+1.37
PBARCITY	-6.457	+3.267	-1.976	-2.97****
PCPOOR	+1.916	+1.315	+1.456	---
CLR	-0.837	+0.343	-2.438	-0.08
CONSTANT	+1.357	---	---	---

*The value of minus twice the log of the likelihood function is 7880.36.

**The estimated coefficient divided by its estimated standard error is presented in this column. These values can be compared to the Standard Normal Density to determine significance levels.

***The means calculated from the random sample of 9841 observations used in estimation of the model are; percent victimized, .140, PREP, .537, LOSS, 119.671, PBARCITY .536, CLR .209, PCPOOR, .225.

****This elasticity is the partial effect of PBARCITY on the probability of victimization assuming that RELPREP=PREP-PBARCITY is held constant.

can differ in magnitude, we introduce them separately in Table 6. Both variables have negative and significant effects; CLR has a small elasticity; in contrast, the elasticity of PBARCITY is around 3.

The negative effect of PBARCITY on the victimization probability suggests that reporting behavior has an aggregate deterrent effect: in cities where people are perceived to have a higher reporting probability the victimization probability is lower.[23]

Conclusion

Our analysis of the individual observations sample provides strong support for the hypothesis that the perceived victim-specific probability of reporting has a deterrent effect on burglars: households that are more likely to report crimes are less likely to be victimized. Since we believe that reporting is an important example of private law enforcement, we view this finding as an evidence of a deterrent effect of private behavior. Also, this paper provides a microeconomic analysis of crime which improves in some respects over previous studies of the deterrence question. Mainly, we are able to consider the fact that different victims present the offender with different potential payoffs and risks. The variable PREP is a victim-specific deterrent variable as opposed to aggregate measures, such as state averages of the number of imprisoned offenders per offense known used, in former studies. Because the variable LOSS is constructed from information on actual losses, we do not have to assume that illegal payoffs are approximated by,

[23]In this cross-city regression analysis one assumes that different cities can be viewed as different markets due to migration costs faced by offenders. This assumption is made in other studies of deterrence.

say, family income (Ehrlich, 1973). In fact, our loss regression shows that, given income, the illegal return is determined by other variables, particularly type of housing and family structure. Because our predicted deterrent variable PREP is exogeneously given to the burglar, we can identify the causal effect of PREP on victimization and avoid the simultaneity problem encountered in aggregate studies caused by the interaction between aggregate crime and punishment.

In addition to victim-specific deterrence, we find an aggregate deterrent effect, which has public policy implications. If a similar effect exists on other crimes as well, a higher reporting rate reduces the victimization rate because offenders shift from crime to legal activities. This is an important conclusion because reporting can be influenced by public policy. Our results suggest that reporting is strongly influenced by the potential of recovery; in particular, the availability of insurance has a very significant effect. Obviously, the tax deductions for losses from theft are a policy variable which can be used to increase reporting. Also, appropriate policy measures can reduce the cost of insurance and thereby increase reporting. Although our results about the effect of time-cost on reporting are inconclusive, it is possible that lowering this cost by making reporting easier can enhance the tendency of victims to report. The finding of aggregate deterrence also shows that a community action--if possible--to increase reporting can reduce crime.

References

Bartel, Ann P. "An Analysis of Firm Demand for Protection Against Crime," Journal of Legal Studies, 4 (June 1975), pp. 443-478.

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

Clotfelter, Charles T. "Urban Crime and Household Protective Measures," Review of Economics and Statistics, 1977, pp. 499-503.

----------. "Private Security and Public Safety," Journal of Urban Economics, 5 (July 1978), pp. 388-402.

Ehrlich, Issac. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," Journal of Political Economy, 81 (1973), pp. 521-565.

Komeasar, Neil K. "A Theoretical and Empirical Study of Victims of Crime," Journal of Legal Studies, 2 (June 1973), pp. 301-321.

APPENDIX PART A

Table of Definitions

NAME	VARIABLE DEFINITION AND REFERENCES
AGE<25	1 if the household head has age below 25; 0 otherwise.
AGE>55	1 if the household head has age above 55; 0 otherwise.
CHILDREN<12	1 if any children live in the residence; 0 otherwise.
CLR	Burglary clearance rates for city -- Calculated by dividing total number of reported offenses by number of clearances. (data from UCR)
INC=0-3K	1 if the household income is in the range from $0 to $3000; 0 otherwise.
INC=3-6K	1 if the household income is in the range from $3000 to $6000; 0 otherwise.
INC=6-10K	1 if the household income is in the range from $6000 to $10000; 0 otherwise. (not used directly in loss estimation - effects thrown on constant)
INC=10-25K	1 if the household income is in the range from $10000 to $25000; 0 otherwise.
INC>25K	1 if the household income is above $25000; 0 otherwise.
IRREG.HOUSING	1 if the residence is not a house, apartment, or flat; 0 otherwise.
LBARCITY	Mean of expected losses for households in city.
LOSS	Expected loss for household -- Calculated as: LOSS = 0*LPR1 + 13*LPR2 + 50*LPR3 + 162*LPR4 + 375*LPR5 + 700*LPR6.

NAME	VARIABLE DEFINITION AND REFERENCES
LOSS=0	1 if no cash or property was taken; 0 otherwise. (not used directly in reportin estimation - effects thrown on constant)
LOSS=1-25	1 if the value of cash and property taken was in the range $1 to $25; 0 otherwise.
LOSS=26-75	1 if the value of cash and property taken was in the range $26 to $75; 0 otherwise.
LOSS=76-250	1 if the value of cash and property taken was in the range $76 to $250; 0 otherwise.
LOSS=251-500	1 if the value of cash and property taken was in the range $250 to $500; 0 otherwise.
LOSS>500	1 if the value of cash and property taken was above $500; 0 otherwise.
LOW.LOSS.DAMAGE	1 if LOSS=1-25 and PROPDAM=1-50 are both 1.
LPR1	Imputed probability of loss=0 for household.
LPR2	Imputed probability of loss=1-25 for household.
LPR3	Imputed probability of loss=25-75 for household.
LPR4	Imputed probability of loss=75-250 for household.
LPR5	Imputed probability of loss=250-500 for household.
LPR6	Imputed probability of loss>500 for household.

NAME	VARIABLE DEFINITION AND REFERENCES
PBARCITY	Mean of imputed probabilities of reporting for households in city.
PCAPTURE	Product calculated by multiplying PBARCITY times CLR for a household.
PCPOOR	Percent of male city population with age from 14 to 29 and income from $0 to $4000. (from 1970 census documents)
PREP	Imputed probability of reporting for household.
PRESENT	1 if a household member was present during the incident; 0 otherwise.
PROPDAM=0	1 if no property was damaged; 0 otherwise. (not used directly in reporting estimation - effects thrown on constant)
PROPDAM=1-50	1 if the cost of property damage was in the range $1 to $50; 0 otherwise.
PROPDAM>50	1 if the cost of property damage was above $50; 0 otherwise.
RACE=BLACK	1 if the household head is black; 0 otherwise.
RATLOSS	Ratio calculated by dividing LOSS by LBARCITY for a household.
RATPREP	Ratio calculated by dividing PREP by PBARCITY for a household.
RECENT	0 if the household head lived at the current residence on April 1, 1970; 1 otherwise.
RELLOSS	Difference calculated by subtracting LBARCITY from LOSS for a household.

NAME	VARIABLE DEFINITION AND REFERENCES
RELPREP	Difference calculated by subtracting PBARCITY from PREP for a household.
RENT	1 if the residence is rented; 0 otherwise.
VICT	1 if the household had a burglary victimization; 0 otherwise.
WITH.INS	1 if the household was insured against thef 0 otherwise.
2.ADULTS	1 if more than one adult and no children live in the residence; 0 otherwise.

APPENDIX PART B
Table 1
Estimated Loss Model

Explanatory Variable	Estimated Coefficient	Estimated Standard Error	Test Statistic
Class LOSS=0			
CONSTANT	1.319	0.080	16.301
INC=0-3K	0.364	0.076	4.754
INC=3-6K	0.207	0.079	2.625
INC=10-25K	-0.199	0.059	-3.384
INC>25K	-0.840	0.100	-8.342
RACE=BLACK	-0.557	0.050	-11.101
AGE<25	-0.010	0.069	-0.146
AGE>55	0.186	0.065	2.861
CHILDREN<12	0.341	0.064	5.292
2.ADULTS	0.101	0.060	1.671
RENT	0.102	0.055	1.848
IRREG.HOUSING	0.941	0.334	2.814
RECENT	-0.204	0.053	-3.823
Class LOSS=1-25			
CONSTANT	0.866	0.086	10.039
INC=0-3K	0.487	0.081	5.975
INC=3-6K	0.344	0.083	4.105
INC=10-25K	-0.119	0.062	-1.892
INC>25K	-0.524	0.103	-5.065
RACE=BLACK	-0.860	0.054	-15.815
AGE<25	-0.085	0.074	-1.143
AGE>55	0.086	0.069	1.249
CHILDREN<12	0.528	0.069	7.642
2.ADULTS	0.243	0.065	3.738
RENT	-0.041	0.058	-0.698
IRREG.HOUSING	1.642	0.332	4.940
RECENT	-0.155	0.056	-2.736
Class LOSS=26-75			
CONSTANT	-0.363	0.108	-3.346
INC=0-3K	0.401	0.098	4.063
INC=3-6K	0.188	0.103	1.815
INC=10-25K	-0.107	0.078	-1.372
INC>25K	-0.277	0.128	-2.157
RACE=BLACK	-0.417	0.066	-6.271
AGE<25	0.072	0.089	0.808
AGE>55	0.022	0.086	0.256
CHILDREN<12	0.462	0.086	5.357
2.ADULTS	0.274	0.082	3.347
RENT	-0.009	0.073	-0.134
IRREG.HOUSING	1.083	0.378	2.863
RECENT	-0.031	0.070	-0.446

APPENDIX PART B
Table 1
(continued)
Estimated Loss Model

Explanatory Variable	Estimated Coefficient	Estimated Standard Error	Test Statistic
Class LOSS=76-250			
CONSTANT	-0.772	0.120	-6.419
INC=0-3K	0.456	0.106	4.280
INC=3-6K	0.288	0.111	2.580
INC=10-25K	-0.158	0.089	-1.777
INC>25K	-0.255	0.148	-1.714
RACE=BLACK	-0.105	0.072	-1.446
AGE<25	0.050	0.100	0.501
AGE>55	0.182	0.094	1.923
CHILDREN<12	0.216	0.094	2.299
2.ADULTS	0.069	0.089	0.784
RENT	0.129	0.082	1.562
IRREG.HOUSING	1.057	0.395	2.672
RECENT	-0.093	0.078	-1.183
Class LOSS=251-500			
CONSTANT	-0.816	0.127	-6.406
INC=0-3K	0.136	0.116	1.173
INC=3-6K	0.167	0.118	1.412
INC=10-25K	-0.232	0.094	-2.470
INC>25K	-0.413	0.163	-2.535
RACE=BLACK	-0.003	0.077	-0.042
AGE<25	0.048	0.107	0.448
AGE>55	0.083	0.102	0.808
CHILDREN<12	0.121	0.100	1.212
2.ADULTS	0.015	0.095	0.163
RENT	0.077	0.088	0.878
IRREG.HOUSING	1.125	0.413	2.723
RECENT	-0.095	0.084	-1.124

[1]The value of minus twice the log of the likelihood function was 64080.1.

Interpretation of the implications of estimated coefficients in a multinomial LOGIT model requires some care. Let β_{ij} denote the coefficient of the ith independent variable, x_i, for the jth classification of the dependent variable. There are six values for j, the six loss ranges, in the problem at hand. Recall that no coefficients are estimated for the last range due to the fact that probabilities must sum to one. We denote the last state J.

The fact that the probabilities must sum to 1 and that we have normalized on state J causes a problem in the interpretation of individual coefficients. Although if a coefficient β_{ij} is positive then increasing x_i does increase the probability of falling into state j relative to that of falling into state J, it does not follow that increasing x_i will cause an increase in the absolute probability of falling into state j. This is because it may be true that for some value(s) of k, $\beta_{ik} > \beta_{ij}$. The resulting increase in p_k may, through normalization, lead to a decrease in p_j because the sum of the p's must equal one. Therefore, it is helpful to devise a technique which enables us to describe the net effects of changes in each x_i upon the p's.

Because in the particular formulation of the problem which we employed, each independent variable x_i could only attain the value 0 or 1, it was relatively easy to devise a technique to display the effects of the independent variables.

Suppose we wish to display the effect of the kth variable, x_k. We first compute the values $\bar{x}_i$, i = 1,...n, which are the mean values of each x_i for all the sample groups where $x_k = 1$. The values $\bar{x}_1, \ldots, \bar{x}_{Q_j}$ might be thought of as describing the "average" individual for whom $x_k = 1$. We use Q_j to denote the total number of variables for state J. Then, for each j, we compute:

$$p_{j0} = p_j(\bar{x}_1, \ldots, \bar{x}_{k-1}, 0, \bar{x}_{k+1}, \ldots, \bar{x}_{Q_j}) \quad ,$$

and

$$p_{j1} = p_j(\bar{x}_1, \ldots, \bar{x}_{k-1}, 1, \bar{x}_{k+1}, \ldots, \bar{x}_{Q_j}) \quad .$$

The value p_{j0} gives us a "base" probability for the kth variable, and the value $p_{j1} - p_{j0}$ gives us the change in p_j which results from setting $x_k = 1$.

Because the constant term represents a series of traits, no simple interpretation is possible. Though it would be feasible to bifurcate the sample on the basis of the attributes subsummed in the constant and look at the contrasts in the probabilities, this was not done.

It is also possible to estimate the variances of these effects. For each k and j, let

$$D = \begin{bmatrix} \frac{\partial p_{j0}}{\partial \beta_{11}} \cdots \frac{\partial p_{j0}}{\partial \beta_{Q_1,1}} \cdots \frac{\partial p_{j0}}{\partial \beta_{Q_{J-1},J-1}} \\ \\ \frac{\partial p_{j1}}{\partial \beta_{11}} \cdots \frac{\partial p_{j1}}{\partial \beta_{Q_1,1}} \cdots \frac{\partial p_{j1}}{\partial \beta_{Q_{J-1},J-1}} \end{bmatrix}$$

Compute

$$V = D\,C\,D^T \, .$$

Then $\mathrm{var}(p_{j0}) = v_{11}$

$\mathrm{var}(p_{j1}) = v_{22}$, and

$\mathrm{cov}(p_{j0}, p_{j1}) = v_{12}$.

Consequently, $\mathrm{var}(p_{j1} - p_{j0}) = v_{11} + v_{22} - 2v_{12}$.

APPENDIX PART B

Table 2

Comparative Effects of Variables Evaluated at Means for the Estimated Loss Model

Explanatory Variable	LOSS=0				LOSS=1-25			
	BASE	CHANGE	S.E.	QUO.	BASE	CHANGE	S.E.	QUO.
CONSTANT	.1467	+.2612	.0064	40.78	.1561	+.1106	.0063	17.46
INC=0-3K	.4262	+.0021	.0107	.1913	.2273	+.0364	.0088	4.143
INC=3-6K	.4253	-.0066	.0113	-.5856	.2343	+.0353	.0094	3.771
INC=10-25K	.4190	-.0210	.0088	-2.380	.2648	+.0156	.0044	3.556
INC>25K	.4173	-.1131	.0152	-7.446	.2774	+.0052	.0140	.3679
RACE=BLACK	.4203	-.0244	.0076	-3.198	.3032	-.0927	.0067	-13.87
AGE<25	.4037	+.0117	.0050	2.340	.2731	-.0211	.0090	-2.339
AGE>55	.4045	+.0296	.0095	3.125	.2806	-.0174	.0050	-3.508
CHILDREN<12	.4093	-.0054	.0084	-.6428	.2286	+.0464	.0081	5.742
2.ADULTS	.4213	-.0256	.0055	-4.639	.2446	+.0282	.0081	3.501
RENT	.3945	+.0224	.0082	2.743	.2698	-.0135	.0041	-3.307
IRREG.HOUSING	.4320	-.0839	.0283	-2.960	.2593	+.1645	.0293	5.612
RECENT	.4273	-.0285	.0079	-3.601	.2698	-.0088	.0068	-1.293

	LOSS=26-75				LOSS=76-250			
	BASE	CHANGE	S.E.	QUO.	BASE	CHANGE	S.E.	QUO.
CONSTANT	.1931	-.0979	.0088	-11.11	.2012	-.1371	.0080	-17.20
INC=0-3K	.0887	+.0070	.0057	1.230	.0665	+.0095	.0051	1.847
INC=3-6K	.0898	-.0007	.0059	-.1107	.0632	+.0071	.0053	1.357
INC=10-25K	.0905	+.0053	.0015	3.541	.0545	+.0032	.0009	3.528
INC>25K	.0899	+.0391	.0053	7.384	.0515	+.0224	.0031	7.269
RACE=BLACK	.0905	+.0080	.0045	1.786	.0532	+.0312	.0020	15.55
AGE<25	.0972	+.0028	.0012	2.331	.0663	+.0019	.0008	2.334
AGE>55	.0915	-.0057	.0016	-3.511	.0613	+.0035	.0042	.8374
CHILDREN<12	.0879	+.0125	.0055	2.272	.0743	-.0145	.0019	-7.527
2.ADULTS	.0854	+.0126	.0053	2.393	.0678	-.0041	.0009	-4.604
RENT	.0990	-.0049	.0015	-3.304	.0638	+.0046	.0038	1.207
IRREG.HOUSING	.0890	-.0050	.0165	-.3012	.0697	-.0071	.0142	-.4973
RECENT	.0892	+.0110	.0020	5.357	.0605	+.0074	.0014	5.369

	LOSS=251-500				LOSS>500			
	BASE	CHANGE	S.E.	QUO.	BASE	CHANGE	S.E.	QUO.
CONSTANT	.1443	-.0928	.0048	-19.13	.1586	-.0440	.0037	-11.77
INC=0-3K	.0654	-.0188	.0026	-7.283	.1258	-.0361	.0047	-7.734
INC=3-6K	.0627	-.0118	.0027	-4.407	.1248	-.0234	.0051	-4.545
INC=10-25K	.0568	-.0099	.0034	-2.947	.1144	+.0068	.0019	3.527
INC>25K	.0543	-.0012	.0069	-.1745	.1097	+.[illegible]477	.0065	7.372
RACE=BLACK	.0422	+.0247	.0016	15.25	.0906	+.0532	.0033	16.14
AGE<25	.0520	+.0015	.0006	2.330	.1077	+.0031	.0013	2.331
AGE>55	.0510	-.0032	.0009	-3.513	.1111	-.0069	.0020	-3.524
CHILDREN<12	.0613	-.0120	.0016	-7.599	.1386	-.0271	.0035	-7.713
2.ADULTS	.0549	-.0033	.0007	-4.595	.1261	-.0077	.0017	-4.625
RENT	.0553	-.0028	.0008	-3.296	.1177	-.0059	.0018	-3.307
IRREG.HOUSING	.0486	+.0003	.0125	.0232	.1014	-.0689	.0105	-6.540
RECENT	.0483	+.0059	.0011	5.327	.1050	+.0129	.0024	5.362

APPENDIX PART 3

Estimated Reporting Model[1]

Category Explanatory Variable	INCOME=0-3K	INCOME=3-6K	INCOME=6-10K	INCOME=10-25K	INCOME>25K
LOW.LOSS,DAMAGE	0.448 (1.952)	0.647 (2.826)	0.593 (3.431)	0.538 (3.389)	---
PROPDAM=1-50	0.304 (2.680)	---	0.288 (3.230)	0.480 (6.038)	0.537 (2.635)
PROPDAM>50	0.839 (3.258)	1.296 (4.050)	1.390 (6.921)	1.361 (8.168)	1.029 (2.887)
LOSS=1-25	0.344 (3.085)	---	---	---	---
LOSS=26-75	1.434 (9.696)	0.986 (6.033)	1.623 (13.182)	1.123 (10.705)	1.054 (4.471)
LOSS=76-250	1.691 (10.248)	2.124 (10.027)	2.031 (13.042)	1.723 (12.162)	1.735 (5.290)
LOSS=251-500	2.311 (10.147)	1.863 (8.564)	2.316 (13.558)	2.020 (11.801)	1.557 (4.576)
LOSS>500	2.412 (12.957)	2.349 (11.478)	2.682 (18.225)	3.239 (19.447)	2.746 (8.360)
WITH.INS	1.539 (6.896)	1.303 (5.764)	0.982 (8.169)	0.985 (11.955)	0.882 (5.037)
AGE<25	-0.273 (-2.954)	---	-0.164 (-1.721)	-0.205 (-1.695)	---
AGE>54	---	---	---	0.199 (2.288)	---
SOMEONE.PRESENT	0.757 (6.721)	0.645 (5.010)	0.779 (7.755)	0.908 (9.558)	1.241 (4.817)
RENT	---	-0.344 (-3.149)	-0.162 (-2.098)	---	---
CONSTANT	-1.053	-0.701	-0.901	-1.131	-1.152
-2*LN(L)	3446.23	2520.12	4882.57	5900.74	939.273

[1]These estimates are for 14 of the cities surveyed in 1973 and 1974.

Attachment

The estimated loss and reporting models are used to impute LOSS and PREP to each household. However, due to technical reasons, the estimated reporting model includes three variables which are unobservable for unburglarized households. These variables are property damage, the presence of a household member during a burglary, and insurance coverage (only burglarized households were asked about insurance). Instead of predicting these variables for each household by a regression model, we adopted a cross-tabulation procedure because it is technically easier and the scope of this study does not warrant estimating a formal model for each of these variables.

We divided property damage into three ranges and this, along with the dichotomous variables insurance and presence, created twelve cells in the cross tabulation. We then defined unique household "types" using the following observable characteristics for victimized households:[1] income, age of head, family type, race, rent or own, recent and regular or irregular housing.[2] We next calculated maximum likelihood estimates of the joint probability density for those twelve cells for each unique household type based on our sample of burglarized households.[3] Estimated probabilities in this density

[1] There are 49,460 usable burglary incidents in the entire sample of 39 cities collected between 1972 and 1975.

[2] This generates a potential number of unique household types of 720.

[3] There were 465 unique household types which were present among the 49,460 households reporting a burglary incident.

are denoted $P(C_j|H_i)$, $j = 1,\ldots, 12$, where H_i is the vector of observable characteristics for the unique household type and C_j is the jth configuration of a priori unobservable aspects of the hypothetical burglary.

Given a unique household type, the loss structure of the hypothetical burglary is determined from the loss model. We denote this distribution $P(L_k|H_i)$, $k = 1,\ldots, 6$. The unobservable characteristics of the hypothetical burglary are fixed when a specific cell of the grid for this unique household type is selected. Given a household type and cell, the reporting model is employed to produce six probabilities of reporting which depend on which of the six loss class variables is activated. We denote this as $P(\text{Report}|L_k, C_j, H_i)$. This matrix of information is available for each unique household type and the following calculation was performed to impute a probability that a hypothetical burglary of this unique household type would be reported:

$$P(\text{Household type i reports}) = P(\text{REPORT}|H_i) = \text{PREP}_i =$$

$$\sum_{j=1}^{12}\left(\sum_{k=1}^{6} P(\text{REPORT}|L_k, C_j, H_i)\cdot P(L_k|H_i)\right)\cdot P(C_j|H_i).$$

A similar calculation was used to generate the expected loss to be imputed to household type i for the hypothetical burglary. Note that this expected loss is the expected revenue to the burglar and does not include any component for property damage. The expected loss was defined to be:

$$LOSS_i = \sum_{k=1}^{6} MIDRANGE(L_k) \cdot P(L_k | H_i).$$

A concordance relating household type to a loss distribution and a probability of reporting was produced. Each household record in the entire sample was read, matched if possible with the appropriate unique household type, H_i, and, using the concordance, given a loss structure probability density, $P(L_k|H_i)$ and a reporting probability ($PREP_i$).

Less than one percent of the households in the entire sample could not be matched with any of the unique household types in the concordance. In addition, no pattern of systematically omitted households configurations emerged.