An application of the simultaneous Tobit model: a study of the determinants of criminal recidivism

Robin C. Sickles, Peter Schmidt, and Ann D. Witte

Economists have developed a model of criminal activity based on rational individual choice which began with the work of Gary Becker in 1968. According to this model, "a person commits an offense if the expected utility to him exceeds the utility he could get by using his time and other resources at other activities" (Becker, [3], p. 176). This model predicts that both higher penalties and better legitimate opportunities can deter crime. Most tests of this model to date have concentrated on sanctions variables in order to determine whether people are deterred by the actions of the criminal justice system.[1] These tests have used aggregate data and, thus, either have omitted measures of legitimate opportunities from their model or used highly questionable proxies (e.g., the median income in an area). Studies using individual data have been non-theoretical and focused on whether an individual returns to crime, a simple variable with only two possible outcomes.

This study proposes to use individual data to examine a more complex measure of recidivism which incorporates the usual return-no return measure as a special case. The use of individual data allows one to explore the effects of legitimate opportunities. The dependent variable considered is the total time sentenced during a follow-up period after release from incarceration (*TTMSEN*). The information contained in this variable tells not only whether the releasee was sentenced for crimes after release, but also indicates the severity of those crimes as re-

flected by the total length of the sentences. Since *TTMSEN* cannot be negative, it has a lower limit of zero. The size of its positive values is of interest for at least two reasons. First, it is of interest because of the high cost of imprisonment. With such costs in mind, an understanding of the factors which determine the time an individual is sentenced should be of more than academic interest. Second, the information on time sentenced is interesting because it may be considered a measure of the seriousness of post-release criminal activity. Although there is a good deal of controversy concerning an appropriate measure of seriousness of criminal activity,[2] time sentenced may be a useful measure of seriousness. First, it correlates highly with other measures of seriousness.[3] Second, it avoids the statistical and conceptual problems associated with the analysis of ordinal or non-continuous scales. Finally, time sentenced has the advantage of reflecting the judge's perception of seriousness. This perception may be influenced by: extenuating circumstances, the results of presentencing investigations, and knowledge of the presence or absence of plea bargaining.

The type of modeling technique related to this measure of recidivism has been in the literature for many years [23]. Recent theoretical advances on the properties of the so-called Tobit estimators also have been made [1]. In this paper the equation determining *TTMSEN* is embedded in a system which determines also an economic measure of the payoff to legal activities, the nominal hourly wage rate for the first job after release (denoted *WAGAR*).[4]

The specification of the wage equation is

Dr. Sickles is assistant professor of economics at George Washington University in Washington, D.C.; Dr. Schmidt is professor of economics at Michigan State University in East Lansing; and Dr. Witte is assistant professor of economics at the University of North Carolina in Chapel Hill.

1. For an excellent review of this work, see Blumstein, Cohen, and Nagin [5]. The original test was made by Ehrlich [8].

2. See Rossi et al. [16] for a review.
3. Sellin and Wolfgang [18] find that two measures of seriousness have high correlation (.875 and .938) with statutory sentence. Further, a recent study finds that the Wolfgang-Sellin index of seriousness has a simple correlation of .68 with the actual time sentenced of men in the North Carolina prison system.
4. See Block and Heineke [4] for a model of criminal activity based on an allocation of time between legal and illegal activities.

s Tobit
ts of

Witte

the sentences.
gative, it has a
of its positive
o reasons. First,
e high cost of
s in mind, an
hich determine
ed should be of
econd, the in-
interesting be-
measure of the
hal activity. Al-
of controversy
ure of serious-
sentenced may
sness. First, it
asures of seri-
statistical and
ted with the
nuous scales.
advantage of
of seriousness.
nced by: ex-
sults of pre-
wledge of the
aining.
related to this
the literature
eoretical ad-
-called Tobit
[1]. In this
TTMSEN is
ermines also
yoff to legal
e rate for the
GAR).⁴
equation is

o measures of
and .938) with
finds that the
simple correla-
of men in the

el of criminal
veen legal and

fairly standard, but the *TTMSEN* variable is included as a right hand side endogenous variable. The estimation of systems of equations with an arbitrary number of limited dependent variables recently has been discussed by Sickles [21] and Sickles and Schmidt [22]. The approach, although similar in several ways to the estimation of simultaneous equations, each of whose dependent variables is truncated is substantially different in several respects [2]. The reader interested in the statistical properties of such models is referred to an earlier paper in this *Journal* (Sickles and Schmidt [22]).

Data

The data used in this study consist of information on the post-release activities of a random sample of 641 men who were in prison in North Carolina in 1969 or 1971. The activities of these men were followed for an average period of 37 months after their release from incarceration.⁵ Although the sample was drawn from a single administrative area, the population from which it was taken is representative of medium and minimum custody prison units throughout North Carolina.⁶ The data are exceptionally

5. A follow-up period of three years is commonly recommended for studies of ex-offenders. See Mulvihill and Turmin [13], p. 549) and U.S. President's Commission ([24], p. 68). The follow-up period ranged from three to 71 months. Follow-ups varied because men in prison in 1969 and 1971 were released at different times. Actual studies have used highly variable follow-ups. Most studies use follow-up of two years or less; many follow-ups are of six months or less. Indeed many studies seem to indicate that follow-up periods as short as one year may be sufficient to determine patterns of post-release activities. The authors chose to use information on all individuals regardless of follow-up in order to increase the sample size. A large sample is important when using techniques such as in this study that rely on asymptotic properties of the estimates. This study has adjusted for the variable follow-up period by considering it as one of the independent variables.
6. The sampling frame contained all men (with the exception noted in footnote) imprisoned in the South Piedmont area of North Carolina in 1969 or 1971 and released prior to June, 1973. This sampling frame was dictated by the original purpose for which data were collected (an evaluation of North Carolina's prisoner work release program) and by the source of funding. Men in prison in 1969 and 1971 were chosen as the frame so that the nature of the program evaluated would be known and describable. The years 1969 and 1971 were chosen so that an adequate follow-up period would be possible (data were collected from July, 1973 to July, 1974). South Piedmont was picked because funding under LEAA's pilot cities program was limited to this area. However, individuals incarcerated in South Piedmont are similar to those incarcerated in medium and minimum

complete in three senses. First, the data contain detailed information on post-release activities, since 71 percent of the sample was interviewed an average of 37 months after release.⁷ Second, criminal records are unusually complete. They were obtained by writing down all areas where there was indication that an individual had lived after release and all areas where there was self-report. An FBI check or North Carolina Department of Correction's records indicated that an individual had contact with criminal justice authorities. Finally, it should be emphasized that this data set is essentially a random sample of all releasees.⁸ Unlike many other data sets, it is not limited to individuals on parole.

Single-Equation analysis of time sentenced

The measure of criminal activity used here is the total length of all prison sentences received by an individual during a specific follow-up time after his release from prison. The finding of a positive value for this variable corresponds to the usual definition of recidivism. However, compared to the usual binary (yes/no) measure of recidivism, this variable contains additional information—the length of time sentenced for those individuals who do receive prison sentences.

The nature of this measure of criminal activity (total length, in months, of all prison sentences received, denoted *TTMSEN*) requires care in its statistical analysis. *TTMSEN* cannot be negative, and a value of zero (no criminal activity) is a common finding. Approximately 67 percent of the individuals in the sample received no prison sentence during the follow-up period, and, therefore, had values of zero for *TTMSEN*. The remaining individuals had values of *TTMSEN* ranging from one to 480 months, with a mean of

custody prison units throughout North Carolina. See Witte [25] for details.
7. Interviews were extremely important as they were the main source of information on post-release employment and provided a useful check on the completeness of post-release criminal histories. Wage rates and other job information obtained in these interviews were checked with employers if such information appeared to be at odds with previous labor market experience.
8. See Witte [27] for a detailed description of sampling and data collection methods. Since the original purpose for which the data were collected was an evaluation of the North Carolina prisoner work release program, the population from which it was drawn excluded men convicted of sex offenses, serious drug offenses, or as public drunks.

47 months. A suitable technique for the analysis of such a variable is the Tobit technique. This method has been applied previously to such variables as the amount that an individual spends on a car during a year, which is another example of a variable which is non-negative and often zero. Although this technique has been discussed recently in the criminological literature (e.g. Palmer and Carlson [15]), it has not to the authors' knowledge been used to actually analyze criminal activity. The Tobit technique treats the positive observations in the same way as a least squares regression does, but it attempts also to account for the sizable number of observations which equal zero, which a least squares technique cannot do. Its chief advantage is that it enables using both the observations for which TTMSEN is zero and those for which it is positive.

The mean of the recidivist measure, TTMSEN, was assumed to be a linear function of: a constant term (CNST); the number of prison convictions prior to the one leading to the sample prison sentence that resulted in imprisonment (PCONRP); age (in months) at release (AAR); age (in months) at the time of the individual's first arrest (AFA); a dummy variable equal to one if the individual's record indicates a serious problem with alcohol or hard drug use, and zero otherwise (ALKY); a dummy variable equal to one for whites, and zero for non-whites (RACE); a dummy variable equal to one if the individual was married at the time of release, and equal to zero if he was not (MS); a dummy variable equal to one if the individual had been on the North Carolina work release program prior to release, and equal to zero if he had not (WR); and the wage on the first job after release from the sample sentence (WAGAR).

The selection of this specification for the TTMSEN equation was based on a substantial amount of preliminary analysis which is not reported here. Although it may be argued that additional explanatory variables should be included in the recidivist model, it was found that the specification outlined above captured the most important economic and social determinants of the process that explained a released offender's decision to commit additional crimes. Some of the other explanatory variables considered were: the number of years of school completed; a dummy variable equal to one if the sample imprisonment was for a felony, and zero if for a misdemeanor; a dummy variable equal to one if the sample conviction was for a crime

against property, and equal to zero otherwise; a dummy variable equal to one if the sample conviction was for a crime against a person, and equal to zero otherwise; the length (in months) of the follow-up period, as well as its squared value; whether or not the sample releasee was supervised; the number of previous rule violations while in prison; and the number of convictions not resulting in imprisonment before the sample sentence. None of these variables was found to be significantly related (at the usual confidence levels) to TTMSEN. The estimation technique employed in the simultaneous modeling was carried out using full information maximum likelihood, which requires a numerical (iterative) maximization of the likelihood function. The computational difficulties encountered were serious enough, even with the rather limited set of explanatory variables, to discourage the authors from enlarging that set by reconsidering variables which were not very significant in the single-equation analysis.

The results of the (single equation) Tobit analysis of TTMSEN, using the explanatory variables listed above, are given in Table 1. Most of the coefficients have signs which agree with the study's a priori expectations.[9] The negative coefficient of WR (which indicates a lower value TTMSEN for the individual on work release, other factors held constant) seems reasonable, since work release is expected to ease a releasee's adjustment to the legal sector of the economy as well as to increase the releasee's labor market skills. The positive coefficient of ALKY is also acceptable, given the destabilizing effects of alcohol and drug abuse and the illegality of drug possession and alcohol abuse. The variables AFA and PCONRP measure the individual's presentence history of criminal behavior. The younger an individual is when he begins criminal activities, and the more previous imprisonments he has had, the higher a value of TTMSEN expected; the negative coefficient of AFA and positive coefficient of PCONRP support this expectation. The negative coefficient of MS also seems acceptable since one can expect a married individual to have a more settled domestic environment.

Table 1

Results of Tobit Analysis of TTMSEN

Variable	Coefficient	T-Ratio
CNST	78.538	2.41
RACE	36.220	2.48
WR	−33.415	−2.38
ALKY	56.739	3.79
MS	−45.175	−3.19
PCONRP	2.974	2.97
AFA	−0.194	−1.84
AAR	−0.182	−2.18
WAGAR	−14.420	−1.47

Value of $\hat{\sigma}^2$ = 10938.958

The significance of RACE is striking in light of the inconclusive previous work of others on the effect of race on recidivism. The study's positive coefficient indicates that whites tend to have higher values of TTMSEN than do non-whites (other factors held constant). This agrees with the earlier findings of Kolodney [10], Guze [9], and Schnur [17], but is contrary to those of Mannering [12], Cook [6], and Palmer and Carlson [15].

The negative coefficient of AAR indicates a higher level of TTMSEN for younger releasees than for older ones. Previous analyses have indicated that the effect of AAR is probably non-linear, but this study was unable to pick up this non-linearity by adding the squared value of AAR to the specification. The results could be attributable to the sample which contained only adult offenders. (The average age in the sample was 32 years.)

Finally, WAGAR is considered. The economic model of crime (Becker [3]; Ehrlich [8]) and previous empirical results (Cook [7]; Witte [27]) indicate that better legitimate economic alternatives (higher WAGAR) should reduce criminal activity. The negative coefficient of WAGAR is consistent with this previous work. Note, however, that its level of significance is rather marginal.

Simultaneous analysis of wage and time sentenced

One problem with the model of the previous section is that it is possible to question the exogeneity of WAGAR. That is, there is reason to believe that an individual's decision to engage in criminal activity (as measured, presum-

ably, by TTMSEN) can affect his wage. This is certainly true if the decision to commit a crime is a time allocation decision (Block and Heineke [4]). Consider the case of an individual who intends to support himself through criminal activity. The expectation is that he reduces the time spent searching for a good legitimate job and thus probably settles for a lower wage. Another obvious possibility is that an individual who intends to commit crimes tends to be transient (to avoid law enforcement agencies), and transient jobs tend not to pay well.

A model which simultaneously determines TTMSEN and WAGAR should, therefore, be specified. The statistical model can be written as follows:

$$y_{1t}^* = \gamma_1 y_{2t} + \delta'_1 x_t + \epsilon_{1t}$$

$$y_{2t} = \gamma_2 y_{1t} + \delta'_2 x_t + \epsilon_{2t}$$

where

$$y_{1t} = \begin{cases} y_{1t}^* & \text{when } y_{1t}^* > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Here x_t is a $K \times 1$ vector of exogeneous variables; δ_1 and δ_2 are $K \times 1$ vectors of coefficients; γ_1 and γ_2 are scalars; y_{1t} and y_{2t} are the observed levels of TTMSEN and WAGAR; and y_{1t}^* is the unobserved variable of which y_{1t} is the observed (truncated) counterpart. The included exogenous variables in the TTMSEN equation are given above. The WAGAR equation expresses the mean of WAGAR as a linear function of the following variables: CNST; RACE; MS; AAR; TTMSEN; the number of years of schooling completed (SG); the last two digits of the year of release (YREL)[10]; and a dummy variable which equalled one if the releasee's typical crime was a crime against a person, and zero otherwise (CRMTP2).

The most direct way to estimate the model is by full information maximum likelihood. This estimator is developed in Sickles and Schmidt [22] and is used in this study. It should be noted that alternative approaches to estimation can be taken, the most obvious being a single equation estimator. Lee [11] has outlined a two-stage least squares estimator for a system

---

10. Several other representatives of the year of release were used (e.g., dummy variables for year of release). This had no significant effect on the other estimates.

Table 2

Results of OLS Analysis of WAGAR

Variable	Coefficient	T-Ratio
CNST	−2.717	−1.05
RACE	.2571	3.15
SG	.0402	2.31
YREL	.0622	1.72
MS	.1322	1.63
AAR	.00051	1.51
CRMTP2	.2423	2.20
TTMSEN	− .00079	−1.11

Value of $\hat{\sigma}^2 = .59174$

Table 3

Results of Simultaneous Analysis of TTMSEN

Variable	Coefficient	T-Ratio
CNST	123.281	2.31
RACE	40.010	2.82
WR	−32.242	−2.39
ALKY	57.822	4.06
MS	−42.178	−3.03
PCONRP	2.878	1.58
AFA	−0.200	−1.91
AAR	−0.169	−2.02
WAGAR	−35.613	−1.57

Value of $\hat{\sigma}_1^2 = 10852.873$

Results of Simultaneous Analysis of WAGAR

Variable	Coefficient	T-Ratio
CNST	−3.3878	−2.66
RACE	0.2466	3.04
SG	0.0399	2.33
YREL	0.0706	3.98
MS	0.1556	1.89
AAR	0.000631	1.82
CRMTP2	0.2184	2.00
TTMSEN	0.000833	0.54

Value of $\hat{\sigma}_2^2 = .58489$

of equations similar to the one used here. This estimator is consistent but inefficient.

Also, this type of simultaneous equations Tobit model has the observed value of the truncated dependent variable (TTMSEN) entering the WAGAR equation instead of its unobserved value. An alternative specification developed by Nelson and Olson [14] would have the unobserved value of TTMSEN entering the WAGAR equation. The justification for this type of specification rests on the assertion that an individual who is basing his legal economic decision in part on the risks of illegal activity reacts (potentially) to what he perceives that risk to be. Perceptions are based on what the releasee's observations concerning sentencing practices are, not on the strength of sentiment toward these practices or some other type of unobservable factor.

The single-equation (OLS) estimates for this equation are given in Table 2. As with the TTMSEN equation, other variables were tried in this equation, but had insignificant coefficients; to save space the study does not report all the other runs that were performed.

The positive coefficients of RACE, SG, and AAR agree with the usual findings in estimation of wage or earnings equations. The positive coefficient of YREL is also reasonable since wages were rising throughout the same period. The authors did not have strong a priori feelings about the signs of the coefficients of MS and CRMTP2. Finally, the negative coefficient of TTMSEN is consistent with the authors' belief that (the expectation of) more criminal activity after release lowers the wage at release.

Table 3 presents the results of the simultaneous analysis of TTMSEN and WAGAR. The results are, not surprisingly, different from the single-equation results given in Tables 1 and 2.

However, most of the changes are small. Among the exogenous variables, none of the coefficients changed much, though the levels of significance changed dramatically for PCONRP in the TTMSEN equation, and for YREL in the WAGAR equation. The coefficients of the right hand side endogenous variables were not as stable. The effect of TTMSEN on WAGAR changed sign, though it is insignificant statistically (at usual levels) in both cases. The effect of WAGAR and TTMSEN changed dramatically, its coefficient changing from −14.4 to −35.6. However, it is still only marginally significant.

Contrary to the arguments presented above, this study has failed to find evidence of an effect of TTMSEN on WAGAR. Rather, the state of the world appears to be recursive, with the wage at release influencing future criminal activity, but not being influenced by it. This result is consistent with the models of Becker [3] and Ehrlich [8]. However, it should be stressed that it is only by performing the simultaneous analysis that this conclusion could be reached.

Summary and conclusions

This study has used the simultaneous Tobit technique to explore the relationship between criminal activity and legitimate opportunity. The

results indicate that an individual participates less in criminal activity if he is married, black, not addicted to alcohol or drugs, has little history of prior criminal activity, and is older when released. In addition, his criminal activity level is lower if he participated in North Carolina's prisoner work release program and found a higher-wage job on release (the study's measure of legitimate opportunity). The wage an individual received on the first job after release was likely to be higher if the individual was white, married, older, was convicted of a crime against a person, and had more education.

The results exhibit some interesting aspects. First, from a statistical point of view, they demonstrate the feasibility of applying the simultaneous Tobit model to a real problem of moderate size. Second, from a modeling point of view, they indicate that the relationship between the criminal activity level and legitimate opportunities may be recursive (legitimate opportunity affects the level of criminal activity, but not the reverse), rather than simultaneous as suggested by Block and Heineke [4]. Third, from a policy perspective the results suggest at least four possible ways to reduce the criminal activity of prison releasees: better legitimate opportunities, more education, treatment of alcohol and drug addiction problems, and participation in programs like work release.

REFERENCES

[1] T. Amemiya, "Regression Analysis When the Dependent Variable is Truncated Normal," Econometrica, 41: 997-1016 (1973).

[2] _____, "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal," Econometrica, 42:999-1012 (1974).

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

[4] M. K. Block and J. M. Heineke, "A Labor Theoretic Analysis of the Criminal Choice," American Economic Review, pp. 314-325 (June 1975).

[5] A. Blumstein, J. Cohen, and D. Nagins, Deterrence and Incapacitation: Estimating the Effect of Criminal Sanction on Crime Rates, Washington, D.C., National Academy of Sciences, 1978.

[6] P. Cook, "The Effect of Legitimate Opportunities on the Probability of Parolee Recidivism," Durham, N.C., Institute of Police Science and Public Affairs, 1973.

[7] _____, "The Correctional Carrot: Better Jobs for Parolees," Policy and Analysis, pp. 11-51 (1975).

[8] I. Ehrlich, "Participation in Illegitimate Activities: A

Theoretical and Empirical Investigation," Journal of Political Economy, pp. 521-67 (1973).

[9] S. B. Guze, "A Study of Recidivism Based Upon a Follow-Up of 217 Consecutive Criminals," Journal of Nervous and Mental Disease, pp. 575-80 (1965).

[10] S. E. Kolodney, "A Study of the Characteristics and Recidivism Experience of California Prisoners," Report of the Select Committee on the Administration of Justice, Parole Board Reform in California: Order Out of Chaos, 1970, Sacramento, CA., Assembly of the State of California.

[11] L. F. Lee, "Multivariate Regression and Simultaneous Equations Models with Some Dependent Variables Truncated," Mimeograph, 1976.

[12] J. W. Mannering, "Significant Characteristics of Recidivists," Probation and Parole Association Journal, pp. 211-27 (1958).

[13] D. J. Mulvihill and M. M. Turmin, Crimes of Violence, Vol. 12, U.S. Government Printing Office, Washington, D.C., 1969.

[14] F. D. Nelson and L. Olson, "Specification and Estimation of a Simultaneous Equation Model with Limited Dependent Variables," International Economic Review, pp. 695-710 (1978).

[15] J. Palmer and P. Carlson, "Problems with the Use of Regression Analysis in Prediction Studies," Journal of Research in Crime and Delinquency, pp. 64-81 (1976).

[16] P. H. Rossi, et al., "The Seriousness of Crimes: Normative Structure and Individual Differences," American Sociological Review, pp. 224-37 (1974).

[17] A. C. Schnur, "Prison Conduct and Recidivism," Journal of Criminal Law and Criminology, pp. 36-42 (1949).

[18] T. Sellin and M. Wolfgang, The Measurement of Delinquency, New York, John Wiley, 1964.

[19] P. M. Service, The Recidivism of Persons Released From Facilities of the North Carolina Department of Correction During January-June, 1968, North Carolina Department of Correction, Raleigh, N.C., 1972.

[20] L. Shaw, "Severity of Crime and Length of Sentence," report to the North Carolina Department of Correction, Raleigh, N.C., 1976.

[21] R. C. Sickles, "Simultaneous Equations Models with Truncated Dependent Variables," unpublished Ph.D. dissertation, Chapel Hill, N.C., University of North Carolina, 1976.

[22] R. C. Sickles and P. Schmidt, "Simultaneous Equations Models with Truncated Dependent Variables: A Simultaneous Tobit Model," Journal of Economics and Business, Vol. 11, No. 21 (1978).

[23] J. Tobin "Estimation of Relationships for Limited Dependent Variables," Econometrica, 26: 24-36 (1958).

[24] United States President's Commission on Law Enforcement and the Administration of Justice, Task Force Reports, Corrections, U.S. Government Printing Office, Washington, D.C., 1967.

[25] A. D. Witte, Work Release in North Carolina: An Evaluation of Its Post-Release Effects, Chapel Hill, N.C., Institute for Research Social Sciences, University of North Carolina, 1975.

[26] A. D. Witte and P. Schmidt, "An Analysis of Recidivism, Using the Truncated Lognormal Distribution," Applied Statistics, 26: 302-11 (1977).

[27] A. D. Witte, "Estimating the Economic Model of Crime with Individual Data," working paper 77-66, Chapel Hill, N.C., Department of Economics; University of North Carolina, 1977.