FORMAL PROCESSING AND FUTURE DELINQUENCY: DEVIANCE AMPLIFICATION AS SELECTION ARTIFACT

DOUGLAS A. SMITH
RAYMOND PATERNOSTER

Does referring a case to juvenile court or diverting it affect a person's future delinquent/criminal behavior? Labeling theory suggests that it does, arguing that formal processing by the juvenile justice system is part of a deviance amplification process that ultimately results in increased criminal/delinquent activity. But critics point out that a higher rate of future offending among those referred to court, often interpreted as evidence supporting the deviance amplification argument, could be nothing more than a selection artifact. Specifically, those referred to juvenile court may have more attributes that are related to future offending than do those who are diverted from the system. Under this scenario, differences between these groups in later offending could simply reflect preexisting differences in criminal propensity. This article discusses approaches for testing the deviance amplification argument against the alternative hypothesis of a selection artifact.

INTRODUCTION

Labeling theorists (Lemert, 1951; Becker, 1963) contend that social reactions to initial or primary deviance may restrict one's ability to maintain a conventional lifestyle. Limitations arise because being labeled may create barriers to legitimate employment or lead to social censure from conventional others. This process, described by Tannenbaum (1938: 19–20) as the "dramatization of evil," increases the likelihood that the labeled person will become more involved in and committed to a deviant line of activity than he or she was before the labeling experience. While considerable debate exists regarding the specific intervening mechanisms that lead from being labeled to secondary deviance (see Paternoster and Iovanni, 1989), a fundamental empirical prediction of labeling theory is that being sanctioned or negatively labeled will *increase*

We thank Dan Nagin for helpful comments during the evolution of this article. We also thank Charles Tittle and Debbie Curran for providing the data used in this article and the Computer Science Center at the University of Maryland for providing computer support for this research.

one's involvement in future deviant conduct—a deviance amplification effect.[1]

The possibility that more severe sanctioning of youthful offenders may increase their future delinquent activity has been a rallying point for certain policy approaches in juvenile justice. Proponents of the labeling perspective have long argued that responsible social policy regarding the problem of juvenile misconduct should be based on "a refusal to dramatize the evil" (Tannenbaum, 1938: 20). If attempts by the juvenile justice system to solve the problem of delinquency/crime appear only to solidify a commitment to additional rule breaking, it is best to do as little as possible (see Klein, 1986). This "doing best by doing nothing" ideology in juvenile justice has been expressed at various times as a policy of "radical nonintervention" (Schur, 1973), "decarceration" (Scull, 1977), and "diversion" (Klein, 1975, 1979).

Assessing the deviance amplification hypothesis involves a deceptively simple question: All else being equal, does the degree of formal processing of juvenile offenders increase their future criminal activity? While there is no shortage of research on this question, there is little agreement regarding what the research shows. A persistent point of uncertainty in evaluating the empirical evidence involves whether variables that impact future offending are in fact equal across groups who are handled differently by the juvenile justice system (see critiques of labeling theory by Hirschi, 1975; Tittle, 1975; Wellford, 1975).

To frame ideas, consider the following equation:

$$R = \theta_1 X_1 + \alpha T + u_1, \tag{1}$$

where R is a measure of recividism, X_1 is a vector of variables that are thought to be associated with recidivism, θ_1 is a conformable coefficient vector, T is a dummy variable coded 1 if the individual is referred to court and 0 if not referred, α is the effect of being referred to court on future offending, and u_1 contains both a random component and unmeasured correlates of recidivism. Our substantive interest in this model is the sign and magnitude of the estimate of the coefficient α. Regardless of the functional form of this equation, if T is uncorrelated with u_1, α will be a consistent estimate of the effect of being referred to court on future offending. Under the hypothesis of deviance amplification this coefficient should be positive.

[1] Deterrence theorists are equally concerned with the importance of social reactions to delinquent behavior. However, those working from a deterrence perspective predict that punishment will *reduce* the sanctioned person's involvement in subsequent delinquency—a specific deterrent effect. Because our focus here is on whether previous evidence supporting the secondary deviance component of labeling theory may be the result of a selection artifact, we will not pursue the specific deterrence argument.

But a positive and statistically significant estimate of α can arise for two very distinct reasons. First, referral to court could cause an increase in future offending—a deviance amplification effect. Second, a statistically significant positive coefficient could emerge because of a positive correlation between the variable indicating whether one is referred to court and the disturbance term in the recidivism equation—a selection artifact. In this second case the dummy variable for referral acts as a proxy for correlates of recidivism that are not included as independent variables in the recidivism equation. The important point is that if the variable indicating whether one is referred to juvenile court is correlated with the disturbance term in equation (1), the estimated effect of being referred to court on future offending will be biased and inconsistent.[2]

Such bias is potentially widespread in the empirical literature. For example, early research in this area often took the form of simple comparisons of the extent of recidivism between groups of persons who were handled differently by the juvenile justice system (see Wilkens, 1969, and Lipton *et al.*, 1975, for a review of this research). But it soon became apparent that comparing differences in measures of future offending across groups who received different court dispositions was an inadequate test of whether more severe juvenile court interventions had any effect on subsequent delinquent behavior. The basic problem is that assignment to treatment groups (diversion vs. referred to juvenile court, for example) is the result of a nonrandom process in which high-risk youth are more likely to receive more severe dispositions. Thus,

---

2 To see this bias in the case of an ordinary least squares regression equation, let recidivism be a function only of whether one is referred to court or not.

$$R = \alpha T + u_1.$$

The least squares solution is α is:

$$\hat{\alpha} = (T'T)^{-1}\mathrm{T'R}.$$

Substituting the equation for recidivism into this equation yields

$$\hat{\alpha} = \alpha + (T'T)^{-1}T'u_1,$$

so that the expected value for $\hat{\alpha}$ is

$$E(\hat{\alpha}) = \alpha + E[(T'T)^{-1}T'u_1].$$

If the covariance between the variable measuring whether one is referred to court (T) and the disturbance term (u_1) is not zero, the second term on the right-hand side of the las;t equation is not equal to zero and thus the estimated value of α is not equal to its true value. Note also that if the covariance between T and u_1 is positive, $\hat{\alpha} > \alpha$.

those individuals assigned more severe sanctions would be more likely to commit new offenses whether or not any relationship existed between juvenile court dispositions and future offending. It is not surprising that research in this tradition often found that the more severe the sanction, the greater the likelihood of future offending—a finding often interpreted as showing support for labeling theory.

Recognizing these potential problems, researchers sought other approaches to assess the effects of juvenile justice interventions on future criminal activity. One strategy was to use a matching design. As an example of this approach, Gold and Williams (1969) examined the effect of police contacts on the subsequent behavior of apprehended youths and a control group matched on sex, race, age, prior offenses, and recency of last offense. They reported that in twenty of thirty-five matched pairs the apprehended youth committed more subsequent offenses (a deviance amplification effect), in ten pairs the apprehended youth committed fewer offenses (a specific deterrent effect), and in five pairs there was no difference. Testifying to the difficulty of matching groups based on several criteria, however, Gold and Williams were only able to match thirty-five of seventy-four youths in their study who had been apprehended.

One way overcome the difficulty of adequate matching in assuring the comparability of treatment groups is to use random assignment of subjects into these groups. In theory, and in large samples, random assignment of subjects can ensure that different treatment groups will be comparable in terms of extraneous variables which may be related to the dependent variable under investigation—subsequent involvement in delinquency. Note that in relation to equation (1), random assignment to referral or diverted status would imply that T is uncorrelated with u_1, and thus the estimate of α will be a consistent estimate of the true impact of being referred to court on future offending.

Unfortunately, it is not always possible to randomly assign subjects to control and experimental conditions, and there are relatively few published reports of randomized experiments on the relationship between juvenile justice interventions and future delinquency (see the review by Farrington, 1983). In addition, there is often a considerable difference between the design of a randomized experiment and the resulting implementation of the experiment (cf. Empey and Erickson, 1972; Sherman and Berk, 1984).

If matching techniques control for too few relevant variables and randomization strategies are either impractical or fall short in practice, an alternative strategy used in much recent research is to obtain statistical control over extraneous variables by using multivariate statistical models. In this approach, the effects of juvenile justice processing on future criminal activity are estimated while other variables that are hypothesized to influence future offending are controlled for.

Some recent examples of research within this tradition are reported by Horwitz and Wasserman (1979), Rausch (1983), Shannon (1980, 1988), and Wooldredge (1988). Horwitz and Wasserman, for example, examine the relationship between severity of juvenile court disposition and the number of subsequent arrests and report that the severity of juvenile court disposition has a marginal labeling effect, leading to a greater number of subsequent arrests. Shannon's research, also based on bivariate and multivariate models, shows that more severe juvenile court sanctions are significantly associated with an increase in individual's future criminal activity. On the other hand, Wooldredge finds mixed results in his study of the relationship between severity of juvenile court dispositions and future offending.

But while these studies do control for some correlates of future offending when estimating the association between juvenile court sanctions and subsequent delinquency, they may overlook others. Thus, the estimated association between juvenile court status and future delinquency could still be capturing the influence of other variables not included in the analysis but nonetheless correlated with both referral status and criminal propensity. Later in this article we discuss several approaches to this issue that use information about the nonrandom nature of referral decisions when estimating the effect of referral to court on future offending. We also show that failure to consider this information can have profound consequences for the conclusions drawn regarding the relationship between sanctions and future offending.

DATA AND VARIABLES

Data used in this analysis are a subset of cases referred in 1979 to the juvenile justice intake division of the Florida Department of Health and Rehabilitative Services (HRS). Since all juvenile complaints in Florida are processed through the intake division of HRS, the data set contains information on initial decisions to refer youth for formal processing or to handle cases in a variety of informal or nonjudicial ways (see Bishop and Frazier, 1988).

Case data were collected from the central planning and research division of HRS for thirty-one counties in the state. Within each county a random sample of about two hundred cases was selected from all cases referred to juvenile justice intake during 1979. In counties where fewer than two hundred cases were referred, all cases were selected for analysis (see Tittle and Curran, 1988). The number of cases by county range from 82 for Madison County to 214 for Dade County and totaled 5,669 for the entire sample. The following analysis restricts itself to only black and

Table 1. Means of Variables Used in Analysis

Variable	Referred to Court (N=1,544)	Not Referred (N=1,636)
Independent variables		
No. of charges	1.22	1.03
Black youth (1=black, 0=white)	0.31	0.22
Male youth (1=male, 0=female)	0.84	0.76
No. of priors	1.94	0.66
Age	15.06	14.40
Currently under supervision[a]	0.15	0.05
In School[a]	0.76	0.83
Lives with biological parents[a]	0.35	0.45
Lives in single parent household[a]	0.38	0.31
Felony[a]	0.58	0.21
Natural log of intake caseload	5.30	5.35
No. alternative/diversion programs	8.03	9.28
Natural log of county crime rate per 100,000	8.63	8.66
Percentage of county population in urban areas	60.65%	61.47%
Outcome measures based on cases with at least one year of follow-up data (N=2,716)		
Any subsequent referral to intake[a]	0.36	0.23
No. of subsequent referrals to intake[b]	1.07	0.54

[a] 1 = yes, 0 = no.
[b] For the pooled sample the mean number of subsequent referrals is .77 with a standard deviation of 2.0.

white youth referred to HRS for felonies or misdemeanors. This reduces the sample to 3,180 cases.[3]

Table 1 lists the variables used in this analysis. Means of the independent variables are presented separately for the samples of referred and diverted cases. The independent variables include individual and case attributes as well as a few measures of county or jurisdictional characteristics. The sample is 80 percent male and 27 percent black and has an average age of 14.7 years. Other individual level variables in the data include number of charges in the current referral, number of prior referrals, and a series of dummy variables indicating whether the youth is currently under juvenile court supervision, in school, and whether the current referral is for a felony or misdemeanor. Two additional dummy variables identify whether the child resides with both biological parents or lives in a single-parent household. The reference category for these two variables is such other family arrangements as one biological and one step-parent.

In addition to these individual variables, aggregate data are available to measure the caseload of each intake unit (defined as

[3] In Florida HRS processes all juvenile complaints, including those not involving any criminal behavior by the youth such as dependency and neglect cases, truancy, and runaways. These cases, as well as such traffic infractions such as DUI and driving without a license, are excluded from the current analysis.

the number of cases referred to each intake office in 1979 divided by the number of intake officers) and the number of alternative diversion programs available to intake officers when deciding whether to refer a case for formal court processing. Additionally, we include information on county crime rates per 100,000 population and the degree of urbanization for each of the thirty-one counties.[4]

Two indicators are used to measure future criminal activity. One is a binary measure of whether a youth has any subsequent referrals to juvenile justice intake for a felony or misdemeanor offense. However, as Farrington (1987) and others have argued, whether a youth commits an additional offense or not is only one measure of the impact of juvenile justice intervention. Thus, the following analysis will also examine the *number* of referrals to juvenile justice intake for felonies and misdemeanors during the follow-up period.

Since cases were sampled from those brought to the intake division during 1979 and because the data on subsequent referrals were collected by county from November 1980 until March 1981, persons are in the follow-up period for different amounts of time. To address this heterogeneity in exposure risk, we include a variable measuring the number of months from the time of the instant referral until the date on which data on subsequent referrals were collected. On average youths were in the follow-up period for 19.9 months.[5] All else being equal, we expect the probability of any subsequent criminal behavior and the frequency of such behavior to increase with the length of the follow-up period.

Finally, it is worth noting that our data are closely related to the data Bishop and Frazier (1988) used in their analysis of racial disparity in juvenile justice processing. Their data contain 54,266 felony and misdemeanor cases referred to juvenile justice intake in Florida from 1979 to 1981. They report (1988: 250) that for this time period 49.4 percent of youth referred to intake for misdemeanors or felonies are recommended for formal processing. For our much smaller sample of 3,180 cases from 1979 the comparable figure is 49 percent. There are other similarities between the two data sets in the demographics of the samples. In their data 28 percent of youth referred to intake are black, 78 percent are male, with an average age of 15 years. The comparable figures for the data used here are 27 percent black, 80 percent male, with an average age of 14.7 years. Thus, while the data used here contain only a small percentage of the cases processed by HRS in 1979, they ap-

4 The data on crime rates and the percentage of persons residing in urban areas are from 1979 county census data. See Tittle and Curran (1988) for additional details.

5 If a youth turned 18 before the end of the data collection period, his/her time in the follow-up period is the number of months from the instant referral until his/her 18th birthday.

pear representative of the larger population of youths referred to juvenile justice intake in Florida in 1979.

MODELS AND FINDINGS

We first report results from a series of equations for future offending that do not use information about the process by which cases are selected for referral to juvenile court. These results are compatible with equation (1) discussed above. In these models a dummy variable indicating whether intake recommends referral to court is included as an independent variable along with several other variables that are thought to be related to recidivism. Results from four equations are reported in Table 2. The first column lists results from a probit model in which the dependent variable is coded as 1 if the youth has any future referrals for a felony or misdemeanor and 0 otherwise. The dependent variable for the remaining three models—ordinary least squares, tobit, and negative binomial regression—is the actual number of subsequent referrals to juvenile court.[6]

Results of these models show that black, male, and older youth have higher recidivism rates. Subsequent offending is also greater for persons with more prior offenses and for those with more charges against them in the instant offense. Recidivism also varies directly with the crime rate of the county in which the individual resides and with the length of time in the follow-up period. But most central to our concern is the finding that while controlling for these and several other variables, those recommended by intake for formal juvenile court processing are significantly more delinquent during the follow-up period.[7] This finding is consistent whether future criminality is measured as a yes/no variable or examined in terms of the number of future referrals.

Such results have often been interpreted as showing support for the labeling position that formal processing by the juvenile justice system leads to increased future criminal conduct (Horwitz and Wasserman, 1979; Meade, 1974; Shannon, 1988; Thornberry, 1971). But this inference depends on the validity of the assumption that the dummy variable indicating whether one is referred to

[6] Since about 70 percent of persons have no future referrals to intake during the follow-up period, the linearity assumption of OLS is problematic. Thus, we estimate two additional models for future offending. The tobit model is a censored regression model and is discussed in Amemiya (1985). The negative binomial regression model is appropriate when the dependent variable is a count of event such as future offenses. We estimate the version of this model discussed in Cameron and Trivedi (1986) as NEGBIN II. This model is an extension of the Poisson regression model and is appropriate when the mean of the dependent variable is not equal to its variance. In these data the mean number of future referrals is .77 with a variance of 4.0.

[7] In our analysis of future offending, to ensure that each person in the sample was followed for at least one year after the instant offense, we excluded those who were 17 years of age or older. This left us with a sample of 2,716.

Table 2. Models of the Association Between Independent Variables and Subsequent Offending, Ignoring Selection (N = 2,716)

Independent Variable	Probit I		OLS II		Tobit III		Negative Binomial Regression[a] IV	
No. of charges	0.335[b]	(5.83)[c]	0.402	(4.98)	1.142	(5.54)	0.408	(3.65)
Black	0.126	(2.00)	0.178	(2.07)	0.548	(2.19)	0.197	(1.94)
Male	0.399	(5.56)	0.510	(5.63)	2.010	(6.66)	0.990	(8.55)
No. of priors	0.137	(10.51)	0.195	(11.59)	0.462	(10.67)	0.189	(7.88)
Age	0.058	(3.36)	0.021	(1.13)	0.202	(3.36)	0.068	(2.90)
Under supervision	0.111	(1.05)	0.057	(0.38)	0.378	(0.94)	−0.013	(−0.08)
In school	−0.028	(−0.40)	−0.084	(−0.86)	−0.125	(−0.44)	−0.047	(−0.43)
Felony	−0.029	(−0.49)	−0.013	(−0.16)	−0.141	(−0.58)	−0.065	(−0.74)
Biological parents	−0.105	(−1.51)	0.022	(0.23)	−0.313	(−1.12)	−0.122	(−1.16)
Single parent	0.053	(0.76)	0.151	(1.58)	0.312	(1.11)	0.137	(1.26)
County crime rate	0.247	(2.88)	0.345	(3.06)	1.246	(3.56)	0.659	(4.65)
Urbanization	−0.001	(−0.06)	0.001	(0.34)	−0.001	(−0.06)	−0.001	(−0.74)
Months in follow-up[d]	0.025	(3.32)	0.040	(4.06)	0.121	(4.04)	0.916	(4.20)
Intake recommendation	0.148	(2.45)	0.175	(2.13)	0.722	(2.95)	0.245	(2.82)
Constant	−4.953		−4.53		−22.58		−11.45	
Sigma/Alpha[e]					4.344		2.902	(17.48)
Log likelihood	−1,441.50		−5,565.51		−3,040.62		−2,802.4	

[a] This model is described in Cameron and Trivedi (1986) as NEGBIN II.
[b] Coefficients are maximum likelihood estimates except for ordinary least squares which are metric coefficients.
[c] Asymptotic t-ratio.
[d] The natural log of months is used for the negative binomial regression model.
[e] The coefficient in this row for the tobit model is the maximum likelihood estimator of sigma; for the negative binomial regression model this coefficient is the variance parameter alpha defined in Cameron and Trivedi (1986).

juvenile court is independent of the residual term in the equation for future offending. We think that in most empirical research on this topic, this inference is problematic.

Suppose, for example, that a youth's future criminal involvement is associated with some variables that are not measured in the data—say, parental or sibling criminality. Further assume that intake officers are sometimes aware of the criminal histories of other family members and that this knowledge increases the chances that intake will recommend referral to juvenile court. Under this or a number of other plausible scenarios, the treatment status of individuals (referred or diverted) is related to unmeasured characteristics (parental or sibling criminality) that in turn influence the dependent variable of interest (future offending). Thus, the variable measuring whether a case is referred to juvenile court is potentially confounded with unmeasured variables that influence future offending.

Viewed from this perspective, bias in estimating the effect of referral to court on future offending results from common omitted variables that influence both the probability of being referred to court and likelihood of future offending. One way to compensate for such bias is to utilize information about the process by which cases are selected for referral to juvenile court. Let this process be represented by the equation

$$T = \theta_2 X_2 + u_2, \tag{2}$$

where T is a dummy variable coded as 1 if the person is referred to court and 0 otherwise, X_2 is a vector of measured variables that influence the probability of referral, θ_2 is a coefficient vector, and u_2 contains both a random component and unmeasured variables associated with the probability of being referred to court.

Introducing an equation for the process by which persons are selected for referral to court makes explicit the fact that referral status itself is an endogenous variable, potentially influenced by both observed (X_2) and unmeasured (u_2) variables. Estimating equation (2) provides additional information that can be used to correct for the bias in estimating the effect of referral to court on future offending that arises from common omitted variables in the referral and recidivism equations. One way to use this additional information involves estimating the equation for referral and recidivism simultaneously and allowing the disturbance terms between these equation to correlate. A second approach involves using the predicted values from the referral equation as an instrumental variable in the recidivism equation. A third approach involves estimating the residuals from the referral equation, conditional on the independent variables in that equation and whether the person is in fact referred to court or diverted. These condi-

Table 3. Probit Models of Intake Recommendation for Formal Processing

Independent Variable	I		II	
No. of charges	.525[a]	(7.62)[b]	0.584	(8.17)
Black	.261	(4.59)	0.235	(3.99)
Male	.093	(1.53)	0.119	(1.88)
No. of priors	.070	(6.51)	0.075	(6.31)
Age	.073	(6.26)	0.062	(5.07)
Under supervision	.480	(5.08)	0.516	(5.23)
In school	−.110	(−1.75)	−0.182	(−2.80)
Felony	.956	(18.69)	1.000	(18.69)
Biological parents	−.045	(−0.73)	−0.042	(−0.64)
Single parent	.185	(2.89)	0.147	(2.21)
County crime rate	−.242	(−3.26)	−0.395	(−4.17)
Urbanization	.006	(4.07)	0.011	(6.73)
Intake caseload	−.265	(−3.97)	−0.571	(−6.75)
Alternative programs	−.011	(−3.95)	−0.018	(−6.18)
County 1			1.550	(8.40)
County 2			0.897	(4.44)
County 3			1.289	(7.47)
County 4			1.384	(4.66)
County 5			−0.901	(−4.28)
County 6			0.845	(5.62)
County 7			−1.097	(−4.38)
Constant	0.952		3.695	
Log L	−1,788.57		−1647.04	
Percentage correct	72.0%		74.6%	
RIOC	.476		.534	

[a] Maximum likelihood probit coefficient.
[b] Asymptotic t-ratio.

tional residuals are then entered as an independent variable in the recidivism equation.

A Model for the Decision to Refer a Case to Juvenile Court

Since the cornerstone of each of these approaches is a model for the process by which cases are selected for referral to juvenile court, we estimate a series of probit equations for the intake officers' decisions to recommend referral to juvenile court. Results from two equations are presented in Table 3.

The first equation reported in Table 3 (I) shows that decisions to recommend formal processing are related to attributes of individual cases as well as characteristics of counties and intake offices. The probability that intake will recommend referral to juvenile court varies directly with the gravity of the offense (felony or misdemeanor), the number of prior offenses, and whether the youth is currently under court supervision. In addition, black and older youth as well as those living in single-parent households are significantly more likely to be referred to court. Intake recommendations for formal processing also vary with the size of intake caseloads and the number of alternative treatment programs in the jurisdiction. Larger caseloads and more alternative programs decrease the probability that a case will be recommended for referral to juvenile court. Finally, referral decisions vary with two county characteristics. Cases brought to intake in more urban counties are more likely to be referred to court, while those in higher crime rate counties are less likely to be referred to court.[8]

Since our data are drawn from thirty-one countries, the county in which a case is processed may have an independent effect on the probability that intake will recommend referral to juvenile court. To assess this possibility, we examined the proportion of between-county variance in the residuals from equation I in Table 3. Using an analysis of variance model, we found that 12.4 percent of the variance in these residuals was between counties. This procedure also identified seven counties in which the observed proportion of referrals differed significantly from the proportion that would be expected based on the results on the first equation in this table. Thus, we estimated the equation again adding seven dummy variables to represent these counties. These results are shown as equation II in Table 3.

These results show that the probability of being referred to juvenile court can vary significantly from one county to the next. In five of these counties, for example, youth brought to intake are significantly more likely to be referred to juvenile court, while in two others they are much less likely to be referred to court. Moreover, these differences are independent of the other fourteen variables in the equation. Adding these seven county dummy variables also improves the fit of the model. The likelihood of equation 3.II is significantly larger than for the equation that does not include these seven variables, and the percentage of variance in the residuals from this equation that lies between counties is reduced from 12.4 to 2.3. Moreover, this model correctly classifies 74.6 percent of cases with respect to these decisions and reduces classification errors relative to predictions based on chance (RIOC) by 53.4 percent (see Loeber and Dishion 1983).

This model not only offers a better statistical fit to the data; in addition, the specification of equation 3.II is more congenial with the realities of juvenile justice decisionmaking. It is likely that intake offices develop a set of decision rules that shape their decisionmaking and that some intake offices will see referral to juvenile court as a solution while others may see it as part of the problem (Cicourel, 1968; Emerson, 1969). Put simply, the collective beliefs of juvenile intake offices may vary from one intake office to the next regarding the utility of referring cases for formal court processing (Cohen and Kluegel, 1978; Cohen, 1975; Bailey and Peterson, 1981). While we do not know why these differences

[8] Because more urban counties tend to have higher crime rates, we estimated this equation deleting the crime-rate variable. When the crime-rate variable is removed, the coefficient on urbanization remains positive and significant. Moreover, when urbanization is removed from the model, the coefficient on the crime-rate variable remains negative and significant.

emerge, it is clear from these data that such differences exist and should be included when modeling the process by which cases are selected for referral to juvenile court. Thus, equation II in Table 3 is used to represent the process by which cases are selected for referral to juvenile court.

A Bivariate Probit Model

One way to utilize information from the process by which persons are selected for referral to court involves simultaneously estimating two probit equations: one for referral and one for recidivism (see Heckman, 1976, 1978; Meng and Schmidt, 1985).[9] This approach loses some information on future offending by creating a dummy variable for recidivism. If we assume that the disturbance terms in these two equations are normally distributed and that their joint distribution is bivariate normal, we can estimate these two probit equations simultaneously using maximum likelihood methods.[10] Results from the recidivism portion of this model are presented in the first column of Table 4.

These results are generally consistent with those from the univariate probit model shown earlier in Table 2. Persons with more charges in the instant offense and who are black, male, or older have a significantly higher probability of subsequent offending. Also, the probability of recidivism is higher for those with prior records and among those who live in counties with higher crime rates. But there is also one major difference between the results from the univariate and bivariate probit models. In the univariate model, the coefficient for whether intake recommends formal processing is .148 with a *t*-ratio of 2.45, which implies that being referred to juvenile court is positively and significantly associated with recidivism. But in the bivariate probit model, whether intake recommends referral to court is not significantly associated with recidivism. Moreover, the estimated coefficient for this variable is negative in sign (−.134).

The difference between these results is consistent with the argument that a selection artifact operates to create the mistaken impression that being referred to juvenile court is criminogenic. Under the hypothesis of a selection artifact, the variable indicating

[9] We thank an anonymous reviewer of an earlier version of this manuscript for suggesting this approach.

[10] It should be noted that in the probit model the dependent variable is an observed realization of a continuous unobserved variable. In the recidivism equation, for example, this unobserved variable might be called individual criminal propensity. We only observe whether a person's criminal propensity is sufficiently great to manifest itself in any criminal behavior. A point to note is that the disturbance term in the probit model is also assumed to be a continuous variable that is not directly observable. Thus, the assumption of bivariate normality applies to the joint distribution of unobservable continuous variables. The same logic can be applied to a model involving the observed number of future offenses by invoking several thresholds along the latent variable rather than a single threshold as in the probit model.

Table 4. Models of the Association Between Independent Variables and Subsequent Offending Utilizing Information on the Selection Process (N = 2,716)

Independent Variable	Bivariate Probit I		BCG/Heckman II		Instrumental Variable Approach: NLTSLS III		Tobit IV		Negative Binomial Regression[a] V	
No. of charges	.371[b]	(6.43)[c]	.470	(5.37)	.453	(5.26)	1.244	(5.59)	.461	(3.83)
Black	.147	(2.29)	.224	(2.52)	.215	(2.42)	.625	(2.42)	.240	(2.32)
Male	.406	(5.61)	.525	(5.76)	.452	(5.26)	2.038	(6.71)	1.004	(8.46)
No. of priors	.143	(13.79)	.206	(11.64)	.204	(11.56)	.477	(10.32)	.201	(8.18)
Age	.065	(4.01)	.032	(1.68)	.034	(1.72)	.228	(3.56)	.079	(3.31)
Under supervision	.134	(1.29)	.142	(0.91)	.092	(0.61)	.436	(1.06)	.008	(0.05)
In school	−.042	(−0.57)	−.110	(−1.12)	−.109	(−1.13)	−.195	(−0.68)	−.093	(−0.81)
Felony	.067	(0.76)	.159	(1.37)	.128	(1.12)	.151	(0.44)	.079	(0.54)
Biological parents	−.112	(−1.60)	.017	(0.19)	.016	(0.17)	−.331	(−1.18)	−.132	(−1.24)
Single parent	.070	(1.00)	.185	(1.89)	.181	(1.86)	.383	(1.34)	.164	(1.47)
County crime rate	.231	(2.67)	.315	(2.76)	.317	(2.77)	1.174	(3.28)	.606	(4.37)
Urbanization	.001	(0.08)	.001	(0.53)	.001	(0.45)	.001	(0.06)	−.001	(−0.47)
Months in follow-up[d]	.025	(3.38)	.041	(4.13)	.048	(4.41)	.138	(4.22)	1.077	(4.37)
Intake recommendation	−.134	(−0.70)	−.323	(−1.27)	−.254	(−0.97)	−.156	(−0.20)	−.189	(−0.53)
Constant	−4.887		−4.414		−4.597		−22.591		−11.576	
Cov (u_1,u_2)[e]	.185	(1.55)	.327	(2.08)						
Sigma/Alpha[f]							4.353		2.925	(17.45)
Log likelihood	−2,886.5		−5,563.3				−3,044.9		−2,804.5	

[a] This model is described in Cameron and Trivedi (1986) as NEGBIN II.
[b] Coefficients for the bivariate probit, tobit, and negative binomial regression models are maximum likelihood estimates. Metric coefficients are reported for the BCG/Heckman and the nonlinear two-stage least squares models.
[c] Asymptotic t-ratio.
[d] The natural log of months is used for the negative binomial regression model.
[e] For the bivariate probit model the coefficient is the estimated correlation between the disturbance terms in the referral and recidivism equations. For the BCG/Heckman model the coefficient is the estimated covariance between these two terms.
[f] The coefficient in this row for the tobit model is the mle of sigma, for the negative binomial regression model this coefficient is the variance parameter alpha defined in Cameron and Trivedi (1986).

whether one is referred to court acts as a proxy for other variables that correlate with both the probability of being referred to court and future offending. If the omitted variables in the referral and recidivism equation are positively correlated as they are in these data (.185), the estimated effect of being referred to court on future offending will be biased upward in favor of the deviance amplification hypothesis. The fact that the estimated coefficient on the referral variable declines from .148 to −.134 when the disturbance terms between the referral and recidivism equations are allowed to correlate is consistent with this expectation.

The Instrumental Variable Approach

Another approach to purge the referral variable of its correlation with unmeasured causes of recidivism is to use an instrumental variable in place of the referral variable in the recidivism equation. Using the instrumental variable approach does not require any assumptions about the joint distribution of the disturbance terms in the referral and recidivism equation (see Heckman and Robb, 1985). All that is required is information on at least one variable which influences the probability of being referred to juvenile court that is also not a predictor of recidivism. An examination of results from the probit model for whether one is referred to juvenile court (Table 3) reveals several such variables. Specifically, the size of the intake office's caseload and the number of alternative diversion programs in the county significantly reduce the probability that intake will recommend referral to court. These variables have no obvious theoretical relationship to whether an individual commits future offenses and are thus useful instrumental variables. Additionally, whether the instant offense is a felony or misdemeanor has a strong effect on the probability of referral to court but no apparent relationship to the number of future offenses. Finally, results from the probit model reveal that the probability of being referred to court varies with the intake office which handles the case. We speculated that these effects reflect the collective attitudes and beliefs of intake officers regarding the utility of referring a case to juvenile court, and see no compelling theoretical reason to think that these variables have a direct causal effect on individual recidivism.

Following arguments in Hausman (1983) and Barnow *et al.* (1980), we form an instrumental variable for whether one is referred to court in the following way. Using the results from the second probit equation reported in Table 3, we calculate the predicted probability that an individual will be referred to juvenile court [$p(T=1|X_2]$. Then, we regress the dummy variable indicating whether one is referred to court on the predicted probability that they will be referred and all of the independent variables in the recidivism equation. The predicted values from this regression equation are, by construction, independent of the disturbance term in the recidivism equation and are used as an instrumental variable for referral when estimating the equations for recidivism.

Results from three instrumental variable models are shown in Table 4. The results under the third column (NLTSLS, for nonlinear two-stage least squares) use a least squares regression model for the second-stage recidivism equation.[11] Results in the fourth and fifth columns use the instrumental variable described above as an independent variable in a tobit and negative binomial regression model for recidivism.

Results from each of these three instrumental variable models show that referral for the current offense has no significant independent effect on recidivism, and in each of these models the point estimate for this effect is negative. This is consistent with the results from the bivariate probit analysis. Thus, when an instrumental variable is used to purge the referral variable of its association with unmeasured correlates of recidivism, the apparent labeling effect reported in Table 2 disappears.

The Model of Barnow, Cain, and Goldberger

Another approach that utilizes information on how cases are selected for referral to court when estimating the effect of being referred to court on future offending is discussed in Barnow *et al.* (1980) and Heckman and Robb (1985). This approach is motivated by the following equation:

$$R = \alpha T + C^* + \epsilon, \tag{3}$$

in which C^* represents true individual criminal propensity, and T is defined again as a dummy variable indicating whether one is referred to court and R is a measure of recidivism.

If it were possible to measure a person's criminal propensity, we could estimate equation (3) and obtain an unbiased estimate of the effect of being referred to juvenile court on future offending. But in practice, this equation cannot be estimated because true criminal propensity is an unobserved variable. Instead, researchers use variables that are related to true criminal propensity to approximate the model represented by equation (3). If we reconsider equation (1):

$$R = \alpha T + \theta_1 X_1 + u_1,$$

we see that the term $\theta_1 X_1$ is a proxy for true criminal propensity in equation (3). If true criminal propensity were completely captured by the independent variables in this equation (X_1), then esti-

[11] The standard errors of coefficients in the second-stage equation are corrected using the method outlined in Green (1990).

mating this model would produce an unbiased estimate of the effect of being referred to court on future offending. But these control variables are related to true criminal propensity (C^*) by

$$C^* = \theta_1 X_1 + u_3, \tag{4}$$

where u_3 contains both a random component and unmeasured correlates of criminal propensity. Thus, in practice, equation (1) becomes, by substitution,

$$R = \alpha T + \theta_1 X_1 + u_1, \tag{5}$$

where $u_1 = u_3 + \epsilon$. This is just another way of saying that the disturbance term in the recidivism equation may contain unmeasured correlates of criminal propensity. To repeat a central theme of this article, if any of these omitted correlates of criminal propensity (such as parental or sibling criminality) are also related to the selection of cases for referral to juvenile court, then the variable measuring whether a person is referred to juvenile court (T) will be correlated with the disturbance term in equation (5). Under these conditions the estimated coefficient on the variable identifying whether a case is referred to juvenile court will be biased. This is essentially an omitted variable bias.

Barnow and his associates (1980) and Heckman (1978) and colleagues (Heckman and Robb, 1985; Heckman and Hotz, 1989) discuss a two-part model that, under certain assumptions, can be used to correct for this type of selectivity bias in estimating the effect of being referred to juvenile court on future offending.[12] The first step in using this approach is to estimate a probit model for the process by which cases are selected for referral to court, as was done in Table 3. The next step is to calculate the expected value of the residual in this equation for each person in the sample, conditional on that person's scores on the independent variables in the referral equation (X_2) and whether they are referred to juvenile court or not [i.e., $E(u_2|X_2, T)$].[13] These conditional residuals provide information on each youth's score on *unmeasured* variables

[12] A primary assumption in using this approach is that the disturbance terms in the recidivism and referral equations have a bivariate normal distribution. The papers by Heckman and Robb (1985) and Heckman and Hotz (1989) are published with discussions that focus on assumptions in these models, and interested readers may find these exchanges of value. We believe that a critical point in using these models is that there is no one generic cure for selectivity bias and that applications of these models must make substantive sense in the context of specific applications. For that reason we will discuss why we believe this model is appropriate in the context of estimating the effect of being referred to juvenile court on future offending.

[13] These conditional residuals are calculated as follows (see Barnow *et al.*, 1980: 54; Heckman, 1978: 938). Let the predicted value for each case in the probit equation for referral to court be $\hat{T}$. If the case is referred to court (i.e., $T=1$) then the $E(u_2|X^2, T=1)$ is equal to $\phi(\hat{T})|\Phi(\hat{T})$, where ϕ and Φ represent respectively the standard normal density and distribution functions. If the

that influence referral decisions. To clarify this, consider two hypothetical cases whose predicted scores from the referral equation are .7, and −.7. Since these are predicted values in a probit equation, higher values are associated with a higher predicted probability that the case will be referred to juvenile court (for these two scores the corresponding probabilities of being referred to juvenile court are .76, and .24, respectively). What can we infer about these cases if both are referred to juvenile court despite their very different predicted probabilities of being referred?

It is important to note that the predicted probability of whether intake will recommend referral to juvenile is based on variables that are available in the data set. But referral decisions may also be influenced by variables, such as parental criminality, that are not contained in the data set and thus cannot be included in the referral equation. It seems reasonable that, over a large number of cases, those whose predicted probability of being referred to court is only .24 but who are in fact referred to court have more of these *unmeasured* variables that increase the probability of referral than persons whose predicted probability of being referred to court is .76. Thus, among cases referred to juvenile court, those with *lower* predicted probabilities of being referred will have larger (i.e., more positive) residuals. Specifically, the $E(u_2|X_2,\ T=1)$ equals 1.29 for the case whose predicted probability of being referred to court is .24. For the case whose predicted probability of being referred to court is .76, the $E(u_2|X_2,\ T=1)$ equals .41. Thus, individuals with larger conditional expected values of the residuals from the probit (selection) equation will, on average, rank higher on unmeasured variables that increase the probability of being referred to juvenile court.[14]

In sum, values of the conditional residuals from the selection equation are intended to capture the heterogeneity among persons in the sample on unmeasured variables that influence the probability of being referred to juvenile court. This heterogeneity, by itself, will not bias the estimate of the effect of being referred to juvenile court on future offending. But bias will exist if some of the unmeasured variables that influence referral decisions are also correlated with recidivism.

---

case is not referred to juvenile court (i.e., $T=0$) then $E(u_2|X_2,\ T=0)$ is equal to $-\phi(\hat{T})/[1-\Phi(\hat{T})]$.

[14] The same result holds if we consider cases not referred to juvenile court (i.e., $T=0$). For example, the $E(u_2|X_2,\ T=0)$ is equal to -1.29 for the case whose predicted probability of being referred to court is .76. For the person whose predicted probability of being referred to court is .24, the $E(u_2|X_2,\ T=0)$ is equal to $-.441$. The smaller of these two values (-1.29) corresponds to the case that had a higher predicted probability of being referred to court but was in fact not referred to court. Thus, this case probably had fewer unmeasured variables that would increase the probability of being referred to court. Hence, among cases not referred to court, larger residual values correspond to cases that rank higher on unmeasured variables associated with more severe intake recommendations.

Thus, the second part of this model involves estimating an equation for future offending which includes the conditional residuals from the referral equation as an independent variable. The equation for recidivism now becomes:

$$R = \alpha T + \theta_1 X_1 + \sigma_{23}[E(u_2|X_2, T)] + \epsilon, \quad (5)$$

where σ_{23} is the estimated covariance between the disturbance terms in the referral and recidivism equations and ϵ is a random error component.[15] A test of the null hypothesis of no selectivity bias is a test of whether σ_{23} is equal to zero. Estimating equation 5 provides a direct test of this and produces a consistent estimate of the effect of being referred to juvenile court on future offending (see Heckman, 1978).[16]

Results from estimating this model are shown in Table 4 in the second column. Two points are worth noting. First, the estimated covariance between the disturbance terms in the referral and recidivism equations is positive (.327) and significant (t = 2.08), which is consistent with the position that unmeasured variables which increase the probability of being referred to court are significantly associated with a greater likelihood of committing future offenses. Second, when this potential source of bias is taken into account, the estimated effect of intake's recommendation for formal processing on future offending is not significant and is also negative in sign (−.327). This finding is consistent with results from the bivariate probit and instrumental variable models and suggests that, in these data, evidence supporting the deviance amplification thesis is artifactual.

---

[15] An area of some concern and controversy involves the identification of the parameter σ_{23}. If all of the variables in the referral equation are also included as independent variables in the recidivism equation, the parameters in the recidivism equation are only identified by the assumption that the joint distribution of the disturbance terms in the referral and recidivism equation error terms is bivariate normal. This is weak identification and rests on an assumption that is not readily testable. A recent discussion of preliminary work on estimators that are more robust to violations of distributional assumptions in selection models can be found in Duncan (1983, 1986). In the current application, the parameters in the recidivism equation are identified by exclusion restrictions; variables in the referral equation that are not in the recidivism equation. This form of identification is stronger to the degree that the exclusion restrictions are valid (see Olsen, 1980, and Heckman and Robb, 1985, for additional discussion on this point). As we noted earlier, there are variables that influence the probability of being referred to court that have no obvious theoretical relationship to recidivism.

[16] While the parameter estimates from equation (5) are consistent, they may be biased in small samples. Some simulation evidence of this is provided by Stoltzenberg and Relles (1990) when the second-stage regression model is based on samples of fifty cases. But since the bias and variance of a consistent estimator decreases as the sample size increases, our sample size of 2,716 cases is worth noting. Additionally, the estimated standard errors for applying OLS to equation (5) are incorrect and are adjusted using the method discussed by Green (1981).

SUMMARY AND DISCUSSION

We have examined the claim that more formal processing by juvenile justice agencies is part of a deviance amplification process that increases future criminal activity. We have outlined an alternative argument suggesting that the positive association between being referred to court and future offending arises because of a selection artifact. This alternative hypothesis reflects the realities of the process by which persons are selected for further court processing and the limitations inherent in nonexperimental data. When youth are brought to juvenile intake, the staff of these agencies makes distinctions between high- and low-risk youth. To some extent these are subjective judgments based on a set of decisions rules formed from experience. Moreover, intake officers are more likely to refer higher-risk youth to juvenile court. The result is that the sample of persons referred to court will contain more high-risk youth who possess a greater likelihood of future delinquent activity. Since we may never be able to fully measure the factors on which youth are selected for referral to juvenile court, there will be heterogeneity between the samples of referred and diverted cases on factors used by intake officers to make referral decisions.

Selectivity bias arise if this heterogeneity is also related to youths' future offending. If intake officers are able to differentiate high-risk from low-risk youth with some degree of accuracy, decisions to refer cases to juvenile court will be positively correlated with unmeasured variables that also increase future offending. Under this scenario, the variable measuring whether a case is referred to court is confounded with unmeasured variables that are themselves causes of future criminal activity. To the extent such confounding exists, models that ignore this type of selection bias will overestimate the true effect of being referred to court on future offending.

Consistent with this expectation, results from a variety of models which assume that selection bias does not exist (Table 2) show that referral to court has a significant positive effect on recidivism. But further analyses which recognize the potential heterogeneity in risk factors between referred and diverted cases (Table 4) reveal that this apparent labeling effect of court referral can instead be attributed to a selection artifact.

Where does this leave us? We think the results reported here raise issues which future tests of deviance amplification should confront. One of these is that serious consideration be given to the possibility that a selection artifact may be responsible for the association between sanctions and future offending in previous analyses of nonexperimental data. Increased attention to this possibility is necessary to make strong inferences about the effects of sanctions on future behavior. In examining whether formal processing by the juvenile justice system increases future offending, the most appropriate null hypothesis is that juvenile justice processing has no causal effect on future offending. This does not mean that we believe this hypothesis is true. It does mean that the burden of proof rests with those who claim that a causal effect exists. Such claims are strengthened to the degree that rival explanations, such as selection bias, can be ruled out. The literature testing the deviance amplification hypothesis has been deficient on this point. We hope that results reported here stimulate future empirical work to correct this weakness.

We also hope that increased attention to possible selection bias in empirical tests of the deviance amplification hypothesis will lead to more conclusive evidence regarding the effects of sanctions on future criminal activity. Progress in this area requires more careful consideration of the assumptions underlying empirical tests. Each of the models we estimated to correct for selectivity bias invoke different assumptions which are either not testable or are matters for theory to resolve. Some of these models depend on distributional assumptions, others on the validity of specific exclusion restrictions. While we think that the specific exclusion restrictions used in the models we estimated make theoretical sense, these exclusion restrictions can never be proven to be true. Thus, results from our analyses remain subject to some degree of uncertainty. And while the results from alternative models to correct for selection bias are quite consistent in showing that referral to court does not lead to increased future offending, additional replications are essential to enhance confidence in this conclusion.

But it cannot be overlooked that models that ignore possible selectivity bias also make strong assumptions. The most critical of these is the assumption that there are no common unmeasured or omitted variables that influence both the probability of referral and the likelihood of future offending. We think our results cast considerable doubt on the validity of this assumption. Moreover, maintaining this assumption when it is not true can lead to substantial bias in the estimated effect of being referred to court on future offending.

In sum, determining whether the positive association between formal processing and future delinquent activity is the result of deviance amplification or a selection artifact is important for both theoretical reasons and from a public policy perspective. This article has focused on some issues in testing these alternative positions and finds no empirical support for the deviance amplification hypothesis. We hope that increased attention to the issues we have discussed and additional research will being us closer to resolving ongoing debates in theory and public policy regarding the effect of sanctions on future offending.

REFERENCES

AMEMIYA, Takeshi (1985) *Advanced Econometrics.* Cambridge, MA: Harvard University Press.

BAILEY, William C., and Ruth D. PETERSON (1981) "Legal Versus Extra-legal Determinants of Juvenile Court Dispositions," 32 *Juvenile and Family Court Journal* 41.

BARNOW, B.S., G. CAIN, and A. GOLDBERGER (1980) "Issues in the Analysis of Selectivity Bias," in E. Stromsdorfer and G. Farkas (eds.), 5 *Evaluation Studies.* San Francisco: Sage Publications.

BECKER, Howard (1963) *Outsiders.* New York: Free Press.

BISHOP, Donna, and Charles FRAZIER (1988) "The Influence of Race in Juvenile Justice Processing," 25 *Journal of Research in Crime and Delinquency* 242.

CAMERON, A. Colin, And P.K. TRIVEDI (1986) "Econometric Models Based on Count Data: Comparison and Applications of Some Estimators and Tests," 1 *Journal of Applied Econometrics* 29.

CICOUREL, Aaron V. (1968) *The Social Organization of Juvenile Justice.* New York: Wiley.

COHEN, Lawrence E. (1975) "Delinquency Dispositions: An Empirical Analysis of Processing Decisions in Three Juvenile Courts." Washington DC: U.S. Department of Justice.

COHEN, Lawrence E., and James KLUGEL (1978) "Determinants of Juvenile Court Dispositions: Ascriptive and Achieved Factors in Two Metropolitan Courts," 43 *American Sociological Review* 162.

DUNCAN, Gregory (1983) "Sample Selection as a Proxy Variable Problem: On the Use and Misuse of Gaussian Selectivity Corrections," *Research in Labor Economics*, Suppl. 2.

——— (1986) "Continuous/Discrete Econometric Models with Unspecified Error Distribution," 32 *Journal of Econometrics* 1.

EMERSON, Robert (1969) *Judging Delinquents: Context and Process in Juvenile Court.* Chicago: Aldine.

EMPEY, Lemar T., and Maynard Erickson (1972) *The Provo Experiment.* Lexington, MA: D. C. Heath.

FARRINGTON, David P. (1983) "Randomized Experiments on Crime and Justice," *in* M. Tonry and N. Norris (eds.), 4 *Crime and Justice: An Annual Review of Research.* Chicago: University of Chicago Press.

——— (1987) "Predicting Individual Crime Rates," in D. Gottfredson and M. Tonry (eds.), 9 *Crime and Justice: An Annual Review of Research.* Chicago: University of Chicago Press.

GOLD, Martin, and Jay R. WILLIAMS (1969) "A National Survey of the Aftermath of Apprehension," 3 *Prospectus* 3.

GREEN, William (1981) "Sample Selection Bias as a Specification Error: Comment," 49 *Econometrica* 795.

——— (1990) *Econometric Analysis.* New York: MacMillan.

HECKMAN, James J. (1976) "Simultaneous Equation Models with Continuous and Discrete Endogenous Variables and Structural Shifts," *in* S. Goldfeld and R. Quandt (eds.), *Studies in Non-Linear Estimation.* Cambridge, MA: Ballinger.

——— (1978) "Dummy Endogenous Variables in a Simultaneous Equation System," 46 *Econometrica* 931.

HECKMAN, James J., and V.J. HOTZ (1989) "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," 84 *Journal of the American Statistical Association* 862.

HECKMAN, James J., and R. ROBB (1985) "Alternative Methods for Evaluating the Impact of Interventions," *in* J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data.* New York: Cambridge University Press.

HIRSCHI, Travis (1975) "Labeling Theory and Juvenile Delinquency: An Assessment of the Evidence," *in* W. Gove (ed.), *The Labelling of Deviance: Evaluating a Perspective.* New York: Halstead Press.

HORWITZ, Allan, and Michael WASSERMAN (1979) "The Effect of Social

Control on Delinquent Behavior: A Longitudinal Test," 12 *Sociological Focus* 52.

KLEIN, Malcolm (1975) "Alternative Dispositions for Juvenile Offenders." Unpublished manuscript.

——— (1979) "Deinstitutionalization and Diversion of Juvenile Offenders: A Litany of Impediments," *in* N. Morris and M. Tonry (eds.), 1 *Crime and Justice: An Annual Review of Research*. Chicago: University of Chicago Press.

——— (1986) "Labeling Theory and Delinquency Policy," 13 *Criminal Justice and Behavior* 47.

LEMERT, Edwin M. (1951) *Social Pathology*. New York: McGraw Hill.

LIPTON, Douglas, Robert MARTINSON, and Judith WILKS (1975) *The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies*. New York: Praeger.

LOEBER, Rolf, and Thomas DISHION (1983) "Early Predictors of Male Delinquency: A Review," 94 *Psychological Bulletin* 68.

Maddala, G. (1983) *Limited-dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.

MEADE, Anthony C. (1974) "The Labeling Approach to Delinquency: State of the Theory as a Function of Method," 53 *Social Forces* 83.

MENG, Chun-Lo, and Peter SCHMIDT (1985) "On the Cost of Partial Observability in the Bivariate Probit Model," 26 *International Economic Review* 71.

OLSEN, R.J. (1980) "A Least Squares Correction for Selectivity Bias," 48 *Econometrica* 1815.

PATERNOSTER, Raymond, and Leeann IOVANNI (1989) "The Labeling Perspective and Delinquency: An Elaboration of the Theory and an Assessment of the Evidence," 6 *Justice Quarterly* 359.

RAUSCH, Sharla (1983) "Court Processing Versus Diversion of Status Offenders: A Test of Deterrence and Labeling Theories," 20 *Journal of Research in Crime and Delinquency* 39.

SCHUR, Edwin M. (1973) *Radical Non-intervention: Rethinking the Delinquency Problem*. Englewood Cliffs, NJ: Prentice-Hall.

SCULL, Andrew T. (1977) *Decarceration, Community Treatment and the Deviant: A Radical View*. Englewood Cliffs, NJ: Prentice-Hall.

SHANNON, Lyle W. (1980) "Assessing the Relationship of Adult Criminal Careers to Juvenile Careers," *in* C. Abt (ed.), *Problems in American Social Policy Research*. Cambridge: Abt.

——— (1988) *Criminal Career Continuity*. New York: Human Sciences Press.

SHERMAN, Lawrence W., and Richard A. BERK (1984) "The Specific Deterrent Effects of Arrest for Domestic Assault," 49 *American Sociological Review* 261.

STOLTZENBERG, Ross M., and D. A. RELLES (1990) "Theory Testing in a World of Constrained Research Design: The Significance of Heckman's Censored Sampling Bias Correction for Nonexperimental Research," 18 *Sociological Methods and Research* 395.

TANNENBAUM, Frank (1938) *Crime and the Community*. Boston: Ginn.

THORNBERRY, Terence P. (1971) "Punishment and Crime: The Effect of Legal Dispositions on Subsequent Criminal Behavior." Ph.D. Thesis, University of Pennsylvania.

TITTLE, Charles R. (1975) "Deterrents or Labeling," 53 *Social Forces* 399.

TITTLE, Charles R., and D. A. CURRAN (1988) "Contingencies for Dispositional Disparities in Juvenile Justice," 67 *Social Forces* 23.

WELLFORD, Charles F. (1975) "Labeling Theory and Crime: An Assessment," 22 *Social Problems* 332.

WILKINS, Leslie T. (1969) *Evaluation of Penal Measures*. New York: Random House.

WOOLDREDGE, John (1988) "Differentiating the Effects of Juvenile Court Sentences on Eliminating Recidivism," 25 *Journal of Research in Crime and Delinquency* 264.