Crime, Time, and Punishment: An Exploration of Selection Bias in Sentencing Research

Marjorie S. Zatz and John Hagan

The sentencing decision reflects the culmination of a long series of processing and, thus, selection decisions, with cases leaving the system at each decision point. Accordingly, the substantive implications of bias due to sample selection are of particular concern for sentencing research. In an effort to assess the existence and manifestations of selection bias, the sentencing decision is modeled for three samples, each of which was selected from different stages of the justice process. Event-history data on felony arrests in the State of California over a 3-year period are used, along with a relatively simple analytic technique which reduces such bias. Results indicate that bias is introduced when censored observations are excluded from the analyses. Also, the effects of certain exogenous variables on sentence length differ, depending upon the selection criteria. Of these, the influence of pleading guilty rather than going to trial is especially interesting. Overall, our findings are consistent with the possibility that selectivity bias has concealed effects of sentence bargaining in some earlier studies.

KEY WORDS: selection bias; sentencing; plea bargaining; event history; censoring.

1. INTRODUCTION

The existence and extensiveness of sentencing disparities based upon social class, race/ethnicity, and gender have been the focus of much research. In addressing such questions, researchers typically sample from the population of persons sentenced to incarceration (e.g., Bullock, 1961; Chiricos and Waldo, 1975; Lotz and Hewitt, 1975; Kelly, 1976; Jacobs, 1978; Zatz, 1984) or convicted (e.g., Gibson, 1978; Spohn et al., 1981;
Brereton and Casper, 1981; Radelet, 1981). Still others examine several processing decisions, with earlier outcomes exogenous for analyses of sentencing outcomes (e.g., Bernstein et al., 1977; LaFree, 1980b; Farnworth and Horan, 1980). While these studies explicitly recognize the existence of several stages in justice processing, the final sample once again reflects only persons sentenced. Approaches such as these exclude cases in which a person committed a criminal act but was not arrested, was arrested but released by the police or prosecutor, or was acquitted or dismissed by the court. All of these persons could be subject to incarceration, yet the selection process excludes them prior to the sentencing decision. In so doing, much variation in sentence severity is lost (Greenberg, 1977; Hopkins, 1977; Reasons, 1977; Klepper et al., 1983).

At each stage, a judgment is made either to continue the case to the next decision point or to release the accused. Decisions of whom to arrest, prosecute, convict, and sentence to prison are based on a variety of factors. The primary influences include the wishes of the complainant or victim (Black, 1970; Green, 1970; Reiss, 1971; Hindelang, 1978; LaFree, 1980a, b; Smith and Visher, 1981), quality and quantity of the evidence (Myers, 1979; Myers and Hagan, 1979), caseload considerations (Eisenstein and Jacob, 1977; Feeley, 1979), and severity of the offense (Burke and Turk, 1975; LaFree, 1980a). Given the diversity of these criteria, it is quite likely that the population of cases selected to continue to the sentencing decision differs considerably from the original population of all offenses or even all arrests. If this is indeed the case, then estimates based on the selected population may well be biased.

Essentially, this is a problem of specification error. The bias results from missing data, where the very data that are missing define the selection process. Somewhat surprisingly, sentencing studies have not addressed this problem. At best, researchers have asserted that their results are generalizable only to the population of persons sentenced, that is, to the selected population. In so doing, results would appear to have external validity. Unfortunately, internal validity can also be problematic if selection problems bias parameter estimates. As Berk and Ray assert "... we are not 'just' saying that poor sampling procedures jeopardize external validity; we are saying that internal validity is threatened as well through the introduction of fundamental specification errors" (1982, p. 353). And, if we do not have internal validity, we cannot have external validity.

The few exceptions include Berk and Ray (1982), Klepper et al. (1983), and Garber et al. (1983). Additional studies acknowledge the problem but do not address it empirically (e.g., Hagan, 1974; Burke and Turk, 1975; Reasons, 1977; La Free, 1980a; Farnworth and Horan, 1980).