Journal of Quantitative Criminology, Vol. 6, No. 4, 1990

Modeling the Suppression Effect of Correctional Programs on Juvenile Delinquency

Syni-An Hwang[1]

Studies of the effects of correctional programs on juvenile delinquency have observed that delinquents exhibit a sharp rise in their arrest rates up to the time of intervention. The drop to a lower rate following intervention has been labeled a suppression effect. A controversy has arisen regarding the nature of the suppression effect; some scholars attribute it to the effectiveness of the correctional programs, while others claim that it is due to a selection artifact. In this study, we examine attempts to model such phenomena and point out that the general terms in the model are not identifiable. Without identifiability, one can construct models that attribute the suppression effect either to the correctional program or to the selection artifact. Some identifiable models are proposed and their associate likelihood functions are used to present a process of model-based analysis to analyze data collected originally by the American Institutes for Research. Discussion of the feasibility of this type of probabilistic modeling approach to criminal justice phenomena is also given.

KEY WORDS: suppression effect; correctional program; identifiability; point process; model-based analysis; Poisson process.

1. INTRODUCTION

In studies of the effects of correctional programs on juvenile delinquency, delinquents have exhibited a sharp rise in their arrest rate (police contacts or arrests per year) up to the time of intervention in the form of a correctional program, and then they drop to a lower rate following release from the correctional program. This drop in the arrest rate is sometimes referred to as the suppression effect. Having observed this suppression effect phenomenon, a number of authors (e.g., see Murray *et al.*, 1978; Murray and Cox, 1979) have concluded that the programs they studied are very effective in reducing juvenile delinquency.

[1]New York State Department of Health, 2 University Place, Room 135, Albany, New York 12203.

Others, however (e.g., McCleary *et al.*, 1979; Maltz *et al.*, 1980; Maltz and Pollock, 1980) have showed that intervention occurred at a time when juveniles were particularly active committing offenses and that, even without the intervention of a correctional program, the individual's arrest rate would have dropped. In other words, the observed phenomenon is just the product of a selection artifact.

The difficulty that arises when one attempts to adjudicate between these arguments is that the data available to examine delinquent behavior do not come from a simple random sample (SRS). The juveniles who were institutionalized and whose data are studied had higher arrest rates compared with others not studied. The goal of this paper is to explore the use of the process model-based approach to study the suppression effect, given such a biased sampling situation. A model for the pre- and postintervention criminal activity of an individual juvenile is proposed, and several models are considered for the intervention scheme (the selection mechanism that decides whether to send an individual to a correctional program after each arrest). The modeling approach is illustrated through the reanalysis of data from the American Institutes for Research (AIR) and the suppression effect is reexamined based on some of the models developed.

2. DATA AND GENERAL MODEL

The AIR data consist of the records of juveniles who were committed either to Unified Delinquency Intervention System (UDIS) institutions or to Illinois Department of Corrections (DOC) institutions. The UDIS sample has 246 subjects, the DOC sample has 317, and 23 subjects have been sent to both the UDIS and the DOC. Before the current term of institutionalization, all subjects had received only supervision or probation as correctional treatment.

For each subject the following information was available: race, date of birth, and arrest history. The arrest history contains the date of each arrest, type of crime, whether the arrest resulted in intervention, length of the intervention, etc.

The average age of the sample subjects at the time of institutionalization (DOC or UDIS) was 15.7, and they served an average of 10.8 months until first parole. Subsequently, they were back on the street for an average of 16.8 months (until they turned 18 or were sent back to an institution).

For the purpose of this study, the rather complex juvenile justice legal process is simplified and represented in the form of a failure time model. Failure, in our context, means reincarceration. It is assumed that an individual's "events" occur according to a random marked point process, i.e., a point process where there is an auxiliary variable, called a mark, associated with each point. In our case, the mark is the associated probability for the decision whether or not to send a juvenile for intervention "treatment." In other words, after each arrest the judge examines an individual's record and then decides whether to send the juvenile for intervention according to some criterion. This is referred to as the "intervention scheme." As it is used here, "judge" is a general term representing all people involved in making probation decisions.

Given the structure of the AIR study data, the complete process is broken down into separate pre- and postintervention processes. The preintervention process can be considered as a point process, say $X(u)$, beginning at time 0 (which corresponds to the individual's twelfth birthday) and an associated stopping rule (intervention scheme) that terminates the process under the condition that intervention occurs before an amount of time, Δ, elapses (corresponding to the eighteenth birthday). Let T be the random variable for the intervention time. The intervention scheme can be specified in terms of $P(T=t|X(u), u\leq t)$, the probability of an individual being sent to an institution at time t given the arrest record up to t.

During the postintervention process, the juveniles are observed after they are released from the institution, say at time R, until they turn 18, until their second intervention occurs, or until the date the study ended (say, C). In other words, we observe the postintervention process only from time R up to $\text{Min}\{\Delta T^*, C\}$ is observed. The postintervention process can be modeled with a new process, $X^*(t^*)$, and a new intervention scheme, specified in terms of $P(T^*=t^*|X^*(u), u\leq t^*)$.

Under the condition that the first intervention occurs on or before each individual's 18th birthday, and assuming independence of the pre- and postintervention processes, we can write the likelihood function for the entire process as

$$f(T; X(u), u\leq T\,|\,T\leq\Delta)f(T^*; X^*(u), R\leq u\leq \text{Min}\{\Delta, T^*, C\}) \quad (1)$$

where $f(X(u), u\leq t)$ is the likelihood function for the process $X(u)$ up to a fixed time t and it does not change with the intervention scheme.

We can rewrite the component for the preintervention process as

$$f(T; X(u), u\leq T\,|\,T\leq\Delta)=\frac{f(X(u), u\leq t)P(T=t, t\leq\Delta|X(u), u\leq t)}{P\{T\leq\Delta\}} \quad (2)$$

$P(T\leq\Delta)$ depends on both the process and the intervention scheme. Moreover, since all of the juveniles in our sample have been sent to intervention, the second term in the numerator of the Eq. (2) is simply $P(T=t|X(u), u\leq t)$. We can decompose the likelihood function for the postintervention process in a similar way. In the following section we consider a specific model for the likelihood function f.

It has been shown previously (Hwang, 1984) that the general terms in Eq. (2) are not identifiable. This means that we can derive the same likelihood function with different sets of parameter. While we may induce identifiability by assuming specific function forms for the terms in Eq. (2), any inference we draw will be model-sensitive. In fact, there may be widely disparate models that are consistent with the observed phenomenon. We believe that this is one of the reasons for the controversy surrounding the interpretation of the suppression effect. Below, we develop a variety of identifiable models that we then use to study the suppression effect.

3. SPECIFIC MODEL

Given the structure of the study and the decomposition derived in Eqs. (1) and (2), we can derive the full process model by studying the arrest and intervention processes separately.

3.1. Arrest Process Models

For each individual in the study we assume that pre- and postintervention arrest processes [$X(t)$ and $X^*(t)$] are homogeneous Poisson processes with rates λ and λ^*, respectively. The value of λ^* is equal to the product of ρ and λ, where ρ is a constant reflecting the ratio of arrest rates after intervention to those before. Our justification of Poisson assumption is as follows. In our data file, only those offenses in which a police contact occurs are recorded. Many offenses are not recorded simply because the juveniles have not been arrested (Boland and Wilson, 1978). Therefore, events in the criminal activity process have been deleted to produce what is known as a thinned arrest process, in which those events that were never recorded are considered to have been deleted. In addition, there are no multiplicities (the offenders cannot be arrested twice simultaneously). Then if (1) the deletion of each point is independent of the others and independent of the offense process, and (2) the process is stationary, we can show that the arrest process approximates a homogeneous Poisson process. Further discussion about the thinning process is given by Cox and Isham (1980), Maltz (1984), and Hwang (1984).

Let preintervention arrests be observed at times $t_1, t_2, \ldots, t_{N(t)}$, until the intervention occurs at $T = t$, where $N(t)$ is the number of the arrests up to time t. Similarly, let $t_1^*, t_2^*, \ldots, t_{N^*(t^*)}^*$ and $N^*(t^*)$ be the corresponding terms in the postintervention arrest process. Then we can write the likelihood function of the pre- and postintervention arrest processes for each subject in the study as

$$f(X(u), u \leq t) = \prod_{t=1}^{N(t)} \lambda\, e^{-\lambda(t_i - t_{i-1})} = \lambda^{N(t)}\, e^{-\lambda t} \tag{3}$$

$$f(X^*(u), 0 \leq u \leq \text{Min}\{\Delta, T^*, C\})$$

$$= \begin{cases} (\lambda\rho)^{N^*(t^*-R)}\, e^{-\lambda\rho(t^*-R)}, & \text{for } T^* = t^* \leq \text{Min}\{\Delta, C\} \\ (\lambda\rho)^{N^*(\Delta-R)}\, e^{-\lambda\rho(\Delta-R)}, & \text{for } \Delta \leq \text{Min}\{T^*, C\} \\ (\lambda\rho)^{N^*(C-R)}\, e^{-\lambda\rho(C-R)}, & \text{for } C \leq \text{Min}\{T^*, \Delta\} \end{cases} \tag{4}$$

3.2. Intervention Models

The construction of the intervention model is much more difficult than the construction of the arrest process model. It is very difficult to develop a realistic intervention scheme, and even harder to compute the related intervention probability because there are many factors that we need to consider. To demonstrate our approach and simplify the computation, we use only the information on the number of arrests in the records and the length of time between arrests, and we neglect the other covariate information. Qualitatively, we might expect the following behavior:

(1) the higher the number of arrests in the records, the higher the probability of intervention;
(2) the shorter the time between the present and the previous arrests, the higher the probability of intervention;
(3) there is a higher probability of being sent for intervention after a delinquent is released from an institution compared to before the first intervention; and
(4) there is a lower probability of being sent to intervention after a period of good behavior.

With these rules in mind, we consider the following five models for the intervention scheme:

(a) For each arrest there is a constant probability P of intervention occurring.
(b) At the ith arrest there is a probability P_i of intervention occurring.
(c) Intervention occurs whenever an arrest succeeds the previous arrest by fewer than τ months. If we assume that τ is a fixed number and let $Z_1, Z_2, \ldots,$ be the corresponding interarrest times (i.e., Z_i = time between $i-1$th and ith arrest), then $T = \text{Min}\{S_n = \sum_{i=1}^{n} Z_i, Z_n \leq \tau, n \geq 1\}$.
(d) This is the same as scheme c except we assume that τ is a random variable (varying from arrest to arrest) with a distribution L.

(e) Intervention occurs at a given arrest if an individual has more than k arrests in the past τ months, where we allow τ to be a random variable (varying from individual to individual).

Similarly, we can represent five intervention schemes for the postintervention process by adding asterisks to the corresponding terms (e.g., P changes to P^* in scheme a, assuming that $P < P^*$, $P_i < P_i^*, \ldots$, etc.). The K-nearest-neighbor intervention scheme (scheme e) says that judges will make their decisions based on both the number of arrests in the record and the length of the interarrest times. This is probably the most appealing of all the schemes. Unfortunately, even with τ fixed, computing the distribution of the time to intervention in scheme e means computing the distribution of the time until a Poisson process produces k events in an interval of length τ, is a formidable task (see Huntington and Naus, 1975). Nonetheless, it is possible to obtain rather good estimates of means and standard deviations of parameters using simulation. More discussion of this model is given by Maltz and Pollock (1980) and Tierney (1983).

We now describe the likelihood function for intervention schemes a–d. The intervention mechanism in each case enters into the likelihood function in Eq. (2) through the second term in the numerator and through the denominator. The terms in the numerator are generally easy to compute. In scheme a we know that $P\{T = t \mid X(u), u \leq t\}$ is given by a geometric distribution, i.e., $P\{T = t \mid X(u), u \leq t\} = (1-P)^{N(t)-1}P$. In scheme b we have

$$P\{T = t \mid X(u), u \leq t\} = \left(\prod_{i=1}^{N(t)-1} (1 - P_i)\right) P_{N(t)} \tag{5}$$

For schemes c and d, $P\{T = t \mid X(u), u \leq t\}$ is an indicator function.

Next we derive the term $P\{T \leq \Delta\}$. According to intervention scheme a, we can view the time to intervention as the time to absorption in a simple two-state Markov process. Thus a straightforward calculation shows that $P\{T \leq \Delta\} = 1 - e^{-\lambda \Delta P}$. For scheme b we condition the total number of events that happened before Δ and then use the law of total probability to get

$$\begin{aligned} P\{T \leq \Delta\} &= 1 - P\{T > \Delta\} \\ &= 1 - \sum_{n=0}^{\infty} P\{T > \Delta \mid N(\Delta) = n\} P\{N(\Delta) = n\} \\ &= 1 - \sum_{n=0}^{\infty} \left\{\prod_{i=0}^{n} (1 - P_i)\right\} \frac{(\Delta\lambda)^n e^{-\lambda\Delta}}{n!} \end{aligned} \tag{6}$$

where $P_0 = 0$.

Scheme c has been discussed by Tierney (1983), who shows that

$$P\{T \leq \Delta\} = \sum_{n=1}^{\infty} \{H^n(\Delta - (n-2)\tau) - e^{-\lambda\tau} H^n(\Delta - (n-1)\tau)\}\, e^{-\lambda\tau(n-1)} \tag{7}$$

where H^n is a gamma distribution with parameters n and $n\lambda$. Finally, we refer to Hwang (1984), for a derivation of the likelihood function for scheme d based on standard renewal theory.

4. FITTING THE MODELS

Since our main purpose is to explore the use of the model-based approach to study the suppression effect, we fit to the simplest model developed in the previous section to the AIR data, assuming that the pre- and postintervention arrest processes are homogeneous Poisson processes with rates λ and λ^*, and the intervention scheme a is appropriate. Moreover, we assume that each delinquent in our data file has the same rate parameters (λ and λ^*). We also make the implicit assumption that all of the criminal acts of an individual are of the same type. Finally, we discuss the possibility of relaxing various assumptions in our models.

4.1. Sample Log-Likelihood Function for the Basic Model

From the results in the previous section, we have that the sample log-likelihood function is as follows:

(a) Preintervention process:

$$
\begin{aligned}
L(N, n_i, t_i; \lambda, P) &= \log \prod_{i=1}^{N} f_i(n_i, t_i; \lambda, P) \\
&= \left(\sum_{i=1}^{N} n_i\right) \log \lambda - \lambda \left(\sum_{i=1}^{N} t_i\right) + \left(\sum_{i=1}^{N} n_i - N\right) \log (1-P) \\
&\quad + N \log P - N \log (1 - e^{-6\lambda P}) \qquad (8)
\end{aligned}
$$

where 6 years is the length between age 12 and age 18 (the potential ages for being a juvenile), and

N = total number of juveniles in the sample

n_i = total number of preintervention arrests for ith juvenile

t_i = time between the ith juvenile's twelfth birthday and the first intervention

(b) Postintervention process:

$$
\begin{aligned}
&L^*(N, n_i^*, t_i^*, N_1, N_2, R_i, \Delta_i; \lambda^*, P^*) \\
&= \log \prod_{i=1}^{N} f_i^*(n_i^*, t_i^*, R_i, \Delta_i; \lambda^*, P^*) \\
&= \log \left(\prod_{i=1}^{N_1} (\lambda^*)^{n_i^*} e^{-\lambda^*(t_i^* - R_i)} (1-P^*)^{n_i^*-1} P^* \right.
\end{aligned}
$$

$$\times \prod_{i=1}^{N_2} (\lambda^*)^{n_i^*} e^{-\lambda^*(\Delta_i - R_i)} (1-P^*)^{n_i^*}$$

$$\left. \times \prod_{i=1}^{N_3} (\lambda^*)^{n_i^*} e^{-\lambda^*(C - R_i)} (1-P^*)^{n_i^*-1} \right)$$

$$= \left(\sum_{i=1}^{N} n_i^*\right) \log \lambda^* - \lambda^* \left(\sum_{i=1}^{N_1} (t_i^* - R_i) + \sum_{i=1}^{N_2} (\Delta_i^* - R_i) + \sum_{i=1}^{N_3} (C - R_i)\right)$$

$$+ \left(\left(\sum_{i=1}^{N} n_i^*\right) \log \lambda^* - N_i\right) \log (1-P^*) + N_1 \log P^* \tag{9}$$

where

- C: Censoring time (the end of the project).
- R_i: Time of the ith juvenile being released from the first intervention.
- Δ_i: Time of the ith juvenile's eighteenth birthday.
- n_i^*: Total number of postintervention arrests for the ith juvenile.
- t_i^*: Time between the ith juvenile's twelfth birthday and the second intervention.
- N_1: Total number of juveniles for which the second intervention occurs ($T < \text{Min}\{\Delta_i, C\}$).
- N_3: Total number of juveniles being censored ($C < \text{Min}\{T_i^*, \Delta_i\}$).
- N_2: $= N - N_1 - N_3$.

4.2. Parameter Estimates

For the postintervention process, we get the maximum-likelihood estimators (MLEs),

$$\hat{\lambda}^* = \frac{\sum_{i=1}^{N} n_i^*}{TP}, \qquad \hat{P}^* = \frac{N_1}{\sum_{i=1}^{N} n_i^*} \tag{10}$$

where TP is the total postintervention follow-up time for the juvenile in the sample, i.e.,

$$TP = \sum_{i=1}^{N_1} (t_i^* - R_i) + \sum_{i=1}^{N_2} (\Delta_i - R_i) + \sum_{i=1}^{N_3} (C - R_i)$$

Note that $\hat{\lambda}^*$ is equal to the total number of arrests occurring during the follow-up period divided by the total follow-up time TP. In other words, $\hat{\lambda}^*$ is equal to the average postintervention arrests per year. Also, $\hat{P}^*$ is equal to the proportion of juveniles in the data sample who have a second intervention.

For the preintervention process, the likelihood equations for $\hat{\lambda}$ and $\hat{P}$ cannot be solved explicitly and we tried several numerical methods to approximate them. These numerical results suggest that $\hat{P}$ is equal to, or at least very close to, zero. Thus the MLE may not be a good estimate of P, the probability that the intervention occurs after each arrest. However, we can still make inferences about λ (or $1-\rho$) using the following approach for estimating (for more discussion in this issue, see Hwang, 1984).

Take the derivative of the preintervention log-likelihood function [Eq. (8)] with respect to λ and P set it equal to zero (i.e., derive the likelihood equation). If P is a constant, then by solving the likelihood equation as a function of λ, we get the following equation:

$$\lambda=\left(\frac{\sum_{i=1}^{N} n_i - N}{\sum_{i=1}^{N} t_i}\right)\frac{1}{1-P} \tag{11}$$

The MLEs of λ and P should satisfy the above equation. If $0 \leq \hat{P} \leq \hat{P}^*$ is assumed, then

$$\left(\frac{\sum_{i=1}^{N} n_i - N}{\sum_{i=1}^{N} t_i}\right) \leq \hat{\lambda} \leq \left(\frac{\sum_{i=1}^{N} n_i - N}{\sum_{i=1}^{N} t_i}\right)\left(\frac{1}{1-\hat{P}^*}\right) \tag{12}$$

Our main purpose is to examine the suppression effect, which has been defined by Murray and Cox (1979) as

$$\frac{(\text{preintervention arrest rate}) - (\text{postintervention arrest rate})}{\text{preintervention arrest rate}}$$

In our notation, the suppression effect is equal to $1-\rho$ (note that λ^* is equal to $\lambda\rho$). Thus we rewrite Eq. (12) in terms of the suppression effect:

$$1-\left(\frac{(\sum_{i=1}^{N} t_i)\,\hat{\lambda}^*}{\sum_{i=1}^{N} n_i - N}\right) \leq 1-\hat{\rho} \leq 1-\left(\frac{(\sum_{i=1}^{N} t_i)\,\hat{\lambda}^*}{\sum_{i=1}^{N} n_i - N}\right)(1-\hat{P}^*) \tag{13}$$

When $\hat{P}^*$ is small, the upper and lower bounds of Eq. (13) are very close to each other. As we discuss later, the suppression effect $(1-\rho)$ changes little with different values of $\hat{P}$.

4.3. Analysis and Results

4.3.1. Basic Analysis

The AIR sample can be divided into three subsamples according to the intervention program a juvenile has been through. We label these three subsamples of juveniles, "UDIS," "DOC," and "COM" (UDIS and DOC combination) and denote the full sample AIR. We summarize the observed results used in the sample log-likelihood function in Table I. Using the data in Table I and Eq. (10), we calculate the MLEs of $\hat{\lambda}^*$ and $\hat{P}^*$ for the four groups in Table II. Finally, by inserting the above results into Eq. (13), we get the upper and lower bounds for the suppression effect in Table III. In each case the two bounds are reasonably close to each other. This implies

Table I. Observed Values from the Data

Observed values	UDIS	DOC	COM	AIR
	Preintervention process			
Total arrests	3195	4347	364	7906
Total years	862.11	1171.70	79.76	2113.57
Number for first intervention	246	317	23	586
	Postintervention process			
Total arrests	908	906	43	1857
Total years	353.61	443.49	19.65	816.75
Number for second intervention	45	89	6	140

Table II. The MLEs and Variances of λ^* and P^*

MLE (variance)	UDIS	DOC	COM	AIR
λ^*	2.57	2.04	2.19	2.27
	(7.3×10^{-3})	(4.6×10^{-3})	(1.1×10^{-1})	(2.8×10^{-3})
P^*	5.0×10^{-2}	9.8×10^{-2}	1.4×10^{-1}	7.5×10^{-2}
	(5.2×10^{-5})	(9.8×10^{-5})	(2.8×10^{-3})	(3.7×10^{-5})

that the suppression effect changes little with any reasonable values of $\hat{P}$ and gives us the ability to derive the inference that is of interest.

4.3.2. Modification of the Basic Model

We consider two modifications of the model: one takes into account the variability of $\hat{P}^*$ and the other adjusted the starting point of the model.

Table III. Estimated Upper and Lower Bounds for the Suppression Effect

Estimate	UDIS	DOC	COM	AIR
	Unadjusted starting-point model			
Lower bound	0.25	0.41	0.49	0.34
Upper bound	0.29	0.46	0.56	0.39
Modified upper bound	0.30	0.48	0.63	0.40
	Adjusted starting-point model			
Lower bound	0.19	0.39	0.47	0.31
Upper bound	0.23	0.45	0.55	0.36
Modified upper bound	0.30	0.46	0.60	0.37

To take into account the variability of $\hat{P}^*$, we let $0 \leq \hat{P} \leq \hat{P}^* + 2\tilde{\sigma}^2_{P^*}$, where $\tilde{\sigma}^2_P$, is the variance estimate of $\hat{P}^*$ which we obtained by inverting the sample information matrix. Using $\hat{P}^* + 2\tilde{\sigma}^2_{P^*}$ as an adjusted upper bound for P, we can find the modified upper bound of the suppression effect by inserting $\hat{P}^* + 2\tilde{\sigma}^2_{P^*}$ for $\hat{P}^*$ in Eq. (13). We give the result, labeled as modified upper bound, in Table III.

The upper and modified upper bounds are very close to each other, except for the COM sample, which has a 12% difference (COM also has a very small sample size; $N = 23$). We have also considered similar modifications to λ and λ^*, i.e.,

$$\hat{\rho}_m = \frac{(\hat{\lambda}^* - 2\tilde{\sigma}_{\lambda^*})}{(\hat{\lambda}^* + 2\tilde{\sigma}_{\lambda^*})}$$

Since the variance estimates of λ and λ^* are relatively small, the modification has little effect on the results.

Next, we consider the effect of adjusting the starting point of the process. When constructing the model, we made the assumption that each individual's preintervention arrest process began at the twelfth birthday. Among the 586 juveniles in the AIR sample, there are 284 individuals who have arrest records prior to their twelfth birthday. In other words, those 284 individuals began their criminal activity process prior to age 12. One way to avoid this problem is to use the minimum of either one's first arrest date or one's twelfth birthday as the starting point but this makes the sample log-likelihood function very complicated and difficult to handle. Moreover, among the 302 juveniles whose first arrest occurs after their twelfth birthday, we can also argue that their arrest process actually began much later than their twelfth birthday. Instead, we define the average interarrest time as

$$\frac{\sum_{i=1}^{N} (i\text{th individual's intervention time} - \text{first arrest date})}{\sum_{i=1}^{N} (\text{number of preintervention arrests for } i\text{th individual}) - N}$$

where N is the sample size. We then compute an estimator of the starting point of the arrest process by subtracting the average interarrest time from the average first arrest age. Based on these estimated starting points listed in Table IV, we repeated the analysis and computed the upper and lower bounds of the suppression effect in Table III (labeled as the adjusted starting point model).

4.3.3. Results

In general, the suppression effect estimated under our model is lower than that suggested by the analyses of Murray and Cox (1979). Even if we use the most conservative of our estimates, the upper bound for the suppression effect, we find that AIR as a group has a 39% suppression effect,

Table IV. Adjusted Starting-Point Estimates

Estimate	UDIS	DOC	COM	AIR
Average first-arrest age (years)	12.03	12.17	12.12	12.11
Average interarrest time (years)	0.30	0.29	0.23	0.29
Average starting age (years)	11.73	11.88	11.89	11.82

which is much lower than the two-thirds reduction claimed by Murray and Cox. Modification to include the variation of $\hat{P}^*$ shows very little effect on the results, i.e., 40 vs 39%. Adjusting for the starting point of the arrest process also shows little effect on the outcome, i.e., 37 vs 39%. The suppression effect is even lower when the lower bound of the suppression effect is used as our estimator (e.g., 34% for the AIR sample).

Also, we found that the DOC group has a much higher reduction in the incidence of offenses than does the UDIS group: 46 vs 29% for the upper bound and 41 vs 25% for the lower bound. Selection biases between the UDIS and the DOC samples do not appear to explain the difference. Quantitatively, the selection biases were not apparent: UDIS juveniles averaged 13 preintervention arrests, while DOC had 13.7. Both cohorts have exactly the same average preintervention interarrest time, 0.27 year. Qualitatively, most judges and probation officers said they tried to send the more incorrigible juveniles to DOC, and the more manageable but committable juveniles to UDIS (Murray and Cox, 1979). An intriguing point is that the percentage of juveniles whose second intervention occurs in the DOC sample is much higher than the UDIS sample: 28 vs 18%. But the average postintervention interarrest time for DOC is longer than the one for UDIS: 0.49 vs 0.39 year. This result contradicts the general rule about the intervention scheme—the shorter the interarrest time, the higher the probability of intervention. One explanation of this conflict is that the DOC cohort might have committed more serious postintervention crimes than the UDIS cohort. Therefore, the second intervention occurs much more quickly for the DOC cohort than for the UDIS cohort. A study of index and nonindex (index offenses have been considered as felonious offenses) crimes looked at separately, however, does not support this interpretation.

For the preintervention process, DOC has a slightly higher average number of arrests than UDIS on both index and nonindex crimes, 8.2 vs 7.8 for the index crimes and 5.5 vs 5.2 for the nonindex crimes (see Table V); however, DOC has lower averages than UDIS on both index and nonindex postintervention arrests, 1.5 vs 1.9 for the index crimes and

Table V. Observed Values for Index and Nonindex Crimes

Observed values	UDIS	DOC	COM	AIR
	Preintervention process			
Index arrest	1914	2612	233	4749
Nonindex arrest	1281	1735	141	3157
Index intervention	175	236	17	428
	Postintervention process			
Index	457	467	23	947
Nonindex arrest	451	439	20	910
Index intervention	30	69	4	103

1.4 vs 1.8 for the nonindex crimes. While these results suggest that there is some weak evidence of the selection bias between DOC and UDIS (DOC sample has higher averages on both index and nonindex preintervention arrests), the results also suggest that the judge might treat the juveniles in the DOC program more severely than the juveniles in the UDIS program. Even if they commit the same crime, the juveniles from the DOC programs are sent back to the programs much more quickly than the UDIS juveniles. In other words, there are treatment biases between the juveniles in the two samples in the postintervention process. The reason for such differential treatment between DOC and UDIS subjects is not clear to us; however, both the selection and the treatment biases are in the opposite direction from the observed effect. If the selection and treatment biases do exist, then the differences in a suppression effect between DOC and UDIS may be even larger than that observed.

4.3.4. Extension of the Model

There are different ways to build a model in a model-based approach. Here we constructed a simple and basic model and then extended this basic model by relaxing some of the assumptions. For example, we might like our model to include the covariate information that reflects the seriousness of the offense. One way to take into account the differential seriousness of offenses in the model is to give a higher weight to a very serious offense and a lower weight to a less serious offense. Using these weights, we could construct a model employing a scoring rule based on the total weight in the arrest process. We note, however, that it is very difficult to determine sensible weights for various offenses (Sellin and Wolfgang, 1964). Instead, we might assume that there exist separate arrest processes for different types of offenses. To simplify our approach, we categorized the offenses into index and nonindex offenses. Generally speaking, we consider index offenses as felonious offenses.

We assume that an individual's index and nonindex arrest processes are Poisson processes with parameters λ_1 and λ_2, and the probabilities that the intervention occurs after an index and nonindex arrests are P_1 and P_2. Let T_1 and T_2 be the intervention time random variables for the index and nonindex arrest processes. T_1 and T_2 will have the exponential distributions with parameters $\lambda_1 P_1$ and $\lambda_2 P_2$. Let T be the intervention random variable, i.e., $T = \min\{T_1, T_2\}$. If the index and nonindex arrest processes are independent, then T has an exponential distribution with parameter $\lambda_1 P_1 + \lambda_2 P_2$. We define the postintervention parameters in a similar fashion (e.g., $P_1^*, P_2^*, \ldots$, etc.). If we further assume that the last preintervention arrest is the one that induced the intervention, then we can derive the pre- and postintervention sample log-likelihood functions similar to Eqs. (8) and (9).

By applying a similar procedure as in Section 4.3, we estimate the parameters in the sample log-likelihood function. For each fixed P_2, we can find the MLEs of λ_1, λ_2, K, where $P_1 = KP_2$. We summarize some of the results in Table VI, including estimates of the suppression effect for index and nonindex crimes, i.e., $\hat{\rho}_1 = \hat{\lambda}_1^*/\hat{\lambda}_1$ and $\hat{\rho}_2 = \hat{\lambda}_2^*/\hat{\lambda}_2$ (see Table VI). We also give the results for the adjusted starting point process in the same table.

The most interesting finding of this analysis is that the reduction of the arrest rates occurs mainly for index offenses. Actually, there is a two- to fourfold difference between index and nonindex suppression effects (see Table VII). This suggests the following interpretation: after a juvenile has been released from his first intervention, he feels that he will have a much greater probability of being sent back to an intervention should he commit another crime. Moreover, if he commits an index crime, the probability of being sent back to an institution will be even higher than if he commits a nonindex crime. As a result, a juvenile may tend to avoid index crimes because of the higher probability of being sent back to an institution. Actually, the reduction that we observe in the arrest rate might have little connection with the correctional programs to which a juvenile has been sent. It may be that the more important factor in reducing juvenile crime is the increased probability of a second intervention, and not the actual type of intervention. The remainder of the results in this analysis parallel those for the basic model, i.e., DOC does better than UDIS in reducing the number of offenses, and the estimated suppression effect is much lower than Murray and Cox's finding.

Another possible extension results from relaxing the assumption that the arrest rate λ is a constant from individual to individual and allowing λ to be a radom variable. Hwang (1984) reports some preliminary studies

Table VI. MLEs for the Index and Nonindex Models

	UDIS	DOC	COM	AIR
		Postintervention		
MLEs				
$\hat{\lambda}_1^*$	1.29	1.05	1.16	1.16
$\hat{\lambda}_2^*$	1.28	0.99	1.03	1.11
$\hat{P}_1^*$	6.56×10^{-2}	1.48×10^{-1}	1.78×10^{-1}	1.09×10^{-1}
$\hat{P}_2^*$	1.26×10^{-8}	7.11×10^{-9}	5.16×10^{-10}	1.23×10^{-10}
		Preintervention		
$P_2=10^{-2}$				
$\hat{\lambda}_1$	2.05	2.06	2.63	2.08
$\hat{\lambda}_2$	1.41	1.41	1.70	1.42
$\hat{K}$	1.54	1.72	1.60	1.64
$P_2=10^{-7}$				
$\hat{\lambda}_1$	2.02	2.03	2.58	2.04
$\hat{\lambda}_2$	1.40	1.41	1.70	1.42
$\hat{K}$	1.71	2.03	1.88	1.88

in which λ has a gamma distribution; the estimated suppression effect under this model is also much lower than Murray and Cox's result.

5. CONCLUSIONS

In this paper, we considered the problem of estimating the effect of correctional programs on juvenile delinquents. We noted that the controversy of whether the observed suppression effect is due to the correctional program or to a selection artifact is a result of the fact that the general terms in our model are not identifiable. As we mentioned in Section 2, without the identifiability, we can construct models that attribute the observed suppression effect either to the correctional program or to the selection artifact.

Table VII. Estimated Suppression Effect for Index and Nonindex Model

	UDIS	DOC	COM	AIR
$P_2=10^{-2}$				
Index crime	0.37	0.49	0.56	0.44
Nonindex crime	9.5×10^{-2}	0.30	0.40	0.23
$P_2=10^{-7}$				
Index crime	0.36	0.48	0.55	0.43
Nonindex crime	9.1×10^{-2}	0.30	0.39	0.22

The other important issue we considered here concerns the fact that the observed data (e.g., the observations in the AIR sample) are not a simple random sample (SRS), i.e., we expect that those juveniles who have been sent to an intervention program will commit more crimes than those who never have been institutionalized. Since we have observed only the juveniles who have been sent to intervention, standard methods of analysis, including regression approaches (which assume the data set is a SRS), do not appear to be reasonable. Instead, we adopted a model-based approach that explicitly considers this selection feature.

To counter the identifiability problem, we constructed a series of identifiable models based on some simple but reasonable general rules. The suppression effect was studied using these identifiable models. To counter the biased-sample problem, we used the conditional probability (conditioning on the fact that all the subjects have been sent to intervention) to derive the likelihood function. The conditional probability used in the likelihood function can adjust for the biased-sample problem in the selection artifact. Other possible selection biases, e.g., that some judges might be more punitive than the others, are not included in the current model. Actually, without further information in the data set we can only assume these are random effects and eliminate them from our analysis. In doing so, we remove most, but not all, of the selection artifact by the conditional probability procedure that was used in the likelihood function.

At this stage, our model is still quite crude. The analyses to date are not strong enough to prove or disprove the contention that the observed suppression effect is due entirely to selection artifact or to correctional program but they do suggest that the observed suppression effect might be a joint effect of the two factors. The analyses also show that the DOC programs outperform the UDIS programs in reducing the crime rate. Moreover, the analyses indicate that an important factor in reducing juvenile crime may be the increased probability of the second intervention, not the actual type of intervention. This result suggests that if the correctional program increases the probability of a second intervention, then the observed suppression effect should be even larger than what we have observed. Actually, the same result implies that if the probability of the first intervention increases, then the overall crime rate among the juveniles might drop to a much lower level than what was observed.

Thus far, we have only the advantages of using the model-based approach. But this approach is not without problems. As the model gets more complicated, we must estimate more parameters and derive more complicated likelihood functions. First, there is no guarantee that extensions to our models will be analytically tractable. Second, the need to estimate more parameters usually means less precision in the analysis. Nonetheless, we find the analyses to date informative, as they make explicit assumptions that are hidden in the more informal approaches adopted by others.

REFERENCES

Blumstein, A. (1974). Seriousness weights in an index of crime. *Am. Soc. Rev.* 39: 854–864.

Boland, B., and Wilson, J. Q. (1978). Age, crime, and punishment. *Public Interest* 24: 22–34.

Cox, D. R., and Hinkley, D. V. (1974). *Theoretical Statistics*, Chapman and Hall, London.

Cox, D. R., and Isham, V. (1980). *Point Processes*, Chapman and Hall, London.

Cox, D. R., and Snell, E. J. (1981). *Applied Statistics—Principles and Examples*, Chapman and Hall, London.

Erickson, M. L., and Jensen, G. F. (1977). Delinquency is still group behavior! Toward revitalizing the group premise in the sociology of deviance. *J. Crim. Law Criminol.* 68: 262–273.

Holden, R. T. (1983). *Failure Time Models for Criminal Recidivism*, Department of Sociology, Yale University, New Haven, Conn.

Huntington, R. J., and Naus, J. T. (1975). A simpler expression for k-th nearest neighbor coincidence probabilities. *Ann. Prob.* 5: 894–896.

Hwang, S. (1984). *Modeling the Suppression Effect of Correctional Programs on Juvenile Delinquency*, Ph.D. thesis, Department of Statistics, Carnegie Mellon University, Pittsburgh, Pa.

Karlin, S., and Taylor, H. M. (1975). *A First Course in Stochastic Process*, Academic Press, New York, San Francisco, London.

Maltz, M. D., and Pollock, S. M. (1980). Artificial inflation of a delinquency rate by a selection artifact. *Operat. Res.* 28: 547–599.

Maltz, M. D., Gordon, A. C., and McDowall, R. (1980). An artifact in pretest-posttest designs: How it can mistakenly make delinquency programs look effective. *Eval. Rev.* 4: 225–240.

Martin, S. E., Sechrest, L. B., and Redner, R. (1981). *New Directions in the Rehabilitation of Criminal Offenders*, National Academy Press, Washington, D.C.

Martinson, R. (1974). "What works? Questions and answers about prison reform. *Public Interest* 35: 22–53.

McCleary, R., Gordon, A. C., McDowell, D., and Maltz, M. D. (1979). How a regression artifact can make any delinquency program look effective. In Sechrest, L. (ed.), *Evaluation Studies Review Annual, 4*, Sage, Beverly Hills, Calif.

Murray, C. A., and Cox, L. A., Jr. (1979). *Beyond Probation, 94*, Sage Library of Social Research, Sage, Beverly Hills, Calif.

Murray, C. A., Thomson, D., and Israel, C. B. (1978). *UDIS: Deinstitutionalizing the Chronic Juvenile Offender*, American Institutes for Research, Washington, D.C.

Sellin, T., and Wolfgang, M. E. (1964). *The Measurement of Delinquency*, Wiley and Sons, New York.

Tierney, L. (1983). A selection artifact in delinquency data revisited. *Operat. Res.* 31(5): 852–865.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci.* 72: 20–22.

Wolfgang, M. E., Figlio, R. M., and Sellin, T. (1972). *Delinquency in a Birth Cohort*, University of Chicago Press, Chicago.