Institute for Research on Poverty
Discussion Paper No. 771-85

Dynamic Models of Criminal Careers

Christopher Flinn
Department of Economics
University of Wisconsin-Madison

February 1985

I am indebted to my colleagues Arthur Goldberger and Charles Manski for many valuable discussions and comments. Detailed discussions with Alfred Blumstein, Jacquelin Cohen, and John Lehoczky were extremely helpful in preparing this revision. Glen Cain and Ariel Pakes also provided helpful comments. This research was partially supported by a grant from the Sloan Foundation to the Institute for Research on Poverty at the University of Wisconsin-Madison.

ABSTRACT

This paper develops behavioral and econometric models for the analysis of criminal careers. By extending the static economic framework for analyzing criminal activity decisions to the multiperiod case, it is possible to generate life-cycle patterns of behavior that are broadly consistent with empirical regularities found in criminological research. These behavioral models also suggest classes of econometric models appropriate for future research in this field. The paper then discusses an econometric model that has been used to investigate the dynamics of labor market attachment, and offers suggestions for the use of models of point processes in research on criminal careers.

Dynamic Models of Criminal Careers

1. DYNAMIC MODELS IN CRIMINAL CAREER RESEARCH

Economists have long been interested in the determinants of criminal activity (e.g., Bentham, 1780), but only in the past few decades has this field of economic applications become something of a growth industry (see, for example, Schmidt and Witte, 1984, and references therein). A number of models of individual decision-making applied to the problem of criminal activity have been proposed, which share several features. First, they all posit rational behavior on the part of individuals, in that, subject to a set of constraints facing the individual, a function characterizing the individual's preferences is maximized. Second, all models recognize that *risk* is an essential component of the decision to engage in criminal activity. In contrast to the purchase of a can of soup, which has a virtually certain level of ultimate satisfaction associated with consumption of the product, the eventual level of satisfaction associated with the decision to undertake criminal activity can only be described probabilistically. All models of criminal activity then must include some method by which the potential outcomes of risky activities may be evaluated. Third, attention is typically restricted to monetary or monetarized yields from criminal activity. In particular, the "psychic" rewards (whether positive or negative) obtained from criminal activity are not explicitly modelled. The aversion that many neoclassical economists have to explaining differentials in behavior through differences in preferences is reflected in the strong and controversial assumption that individuals have identical preferences;[1] all differences in behavior arise through differences in the choice sets individuals face. Finally, the theoretical models which have been formulated are essentially static in nature. These models do not take account of how the criminal and legitimate opportunities expected to prevail in the future affect current criminal activity decisions. Owing to the neglect of these intertemporal considerations, it might be claimed no theory of rational criminal choice has as yet been rigorously formulated.

The report of the Panel on Deterrence and Incapacitation (Blumstein, Cohen, and Nagin, 1978) cited a need for increased behavioral and statistical modelling at the individual level. Section 3 of this paper presents an econometric model of the criminal career that is designed for use with individual-level data. While this econometric model is not explicitly derived from a behavioral model, it does provide a relatively general statistical representation of criminal careers, and the parameters of the model may be interpreted in the context of standard behavioral theories of criminal activity choice. In Section 2 of this paper, I have chosen to develop some simple behavioral models of criminal activity for two reasons. First, I wanted to begin to address the issue of what type of criminal careers these models might generate. Toward this end, I present analytic results when possible, and some limited simulation experiments when analytic results are not available. Second, I wanted to use these behavioral models as a baseline against which some of the statistical models employed in this field could be evaluated. (Some discussion along these lines is contained in Section 3.) Many behavioral assumptions are implicit in the statistical descriptions of criminal careers, and it may be of some value to assess the value of various statistical models not only by their ability to predict behavior (which is typically quite low, see Chaiken and Chaiken, 1981, for example), but also by the degree of correspondence between characteristics of the statistical model and characteristics of a consistent, dynamic model of decision-making and criminal activity. The converse is also obviously true; current empirical knowledge regarding the dynamics of criminal careers must be used as a guide in the construction and evaluation of any theoretical model that purports to describe the criminal activity decision over time.

Structural models of decision-making also serve a related purpose. They are often required for an assessment of the effects of changes in the distributions of rewards and punishments associated with criminal activity on the amount of time spent in these activities. The practical need for structural models was insightfully presented by Marschak (1953), and I will paraphrase his argument here. Let's say we are interested in the development of a model to explain some measure of the degree or intensity of criminal activity, denoted by x. Then, generally speaking, individual differences in x may arise from differences in earnings potentials in legitimate activities, e, background characteristics, b, the distributions of rewards associated with criminal activities, R, and distributions of penalties if apprehended, P. Then we assume there exists a functional relationship between these characteristics $x = x(e, b, R, P; \underset{\sim}{\Omega})$, where $\underset{\sim}{\Omega}$ is the vector of parameters which, in conjunction with the functional form $x(\cdot)$, completely characterizes the relationship between x and the characteristics e, b, R, P. In this case a decision-theoretic model may be of use in guiding our choice of a functional specification of $x(\cdot)$; but once the function is selected the determination of the effects of the exogenous variables on x is simply an empirical matter. The qualitative and quantitative effects of all exogenous variables are contained in the parameter estimates, $\hat{\underset{\sim}{\Omega}}$.

Such an empirically based strategy has at least one advantage over a highly structured approach to the problem. By specifying a flexible functional form for $x(\cdot)$, it is likely that we will be able to capture the observed relationships between the variables well--that is, we will be able to fit the data. We would then be able to assess the effects of changes in the distributions of punishments on the level of criminal activity, for example, comparing $\tilde{x} = x(e, b, R, \tilde{P}; \hat{\underset{\sim}{\Omega}})$ with $\hat{x} = x(e, b, R, P; \hat{\underset{\sim}{\Omega}})$, where $\tilde{P}$ denotes the "new" punishment distribution. This evaluation is straightforward even if x is a highly nonlinear function.

This approach runs into one major problem in practice. To estimate the parameters associated with the exogenous variables, there must exist a sufficient degree of sample variability in these attributes. If we want to assess the effects of the distribution of punishments on criminal activity, the sample members cannot all be subject to the <u>same</u> set of punishment distributions. If all individuals are subject to the same P, at least one element of the parameter vector $\underset{\sim}{\Omega}$ will not be estimable. Even if a few different values of P are present in the sample, thus making it possible to estimate all elements in $\underset{\sim}{\Omega}$, sample variability in P may be so low as to preclude precise estimation of $\underset{\sim}{\Omega}$.

The choice the analyst has then is to ignore the effects of characteristics that vary little or not at all across sample members or to formulate a behavioral model in which these characteristics appear as parameters. For example, assume R and P vary little or not at all in the sample. If we follow the first option, the analyst would estimate a function of the form $x = x_a(e, b; \underset{\sim}{\Theta})$, where x_a is the new functional form and $\underset{\sim}{\Theta}$ the new parameter vector. It is now impossible to say anything concerning the effect of changes in R and P on x. If we follow the second option, we would estimate a function of the form $x = x_b(e, b; R, P, \underset{\sim}{\Delta})$ where we now treat R and P as parametric to the problem, and $\underset{\sim}{\Delta}$ is a vector of other parameters. The functional form of x_b will be derived from an explicit behavioral model. With this approach it will be possible to perform conceptual experiments in which the effects of changes in R and P on x are analyzed. Thus this "structural" approach to modelling behavior is not pursued for reasons of aesthetics; it enables the analyst to perform conceptual experiments which are not possible with models less closely linked with behavioral theory.

2. BEHAVIORAL MODELS OF CRIMINAL BEHAVIOR

In this section we develop a model of the proportion of time allocated to criminal activity in order to analyze how this allocation of time changes as the individual ages and as a function of his previous criminal career. All models are definitionally simplifications of and abstractions from the "real" world. In addition, it may be disquieting to some simply to view criminal behavior as the outcome of a rational calculus. However, if behavior is a manifestation of conscious choice, it seems necessary to posit that individuals make decisions in a way that is consistent with some underlying set of preferences or view of the alternatives facing them. In the models discussed below, individuals are assumed to act rationally.[2] Their preferences and choices are specified in a deliberately limited way. In terms of areas of potential application, these models may be useful in the analysis of commission rates of various types of property crimes. (A list of symbols used in this section is contained in Table 1.)

All three models have a number of common features. Individuals are assumed to be infinitely lived, or, equivalently, to have an unknown length of life T, which is distributed as an exponential random variable. Since the vast majority of individuals seriously engaged in criminal activity are inactive after age 40, the assumption of infinitely lived individuals is not artificial for purposes of analysis.[3] Within the context of these dynamic behavioral models, we will investigate the individual's time allocation decision. The proportion of time spent in crime in period t is denoted θ_t. The total amount of time in each period of life is normalized to one. The remainder of time in each period $(1 - \theta_t)$ is spent in "legitimate" market work, which is compensated at a rate w_t. Leisure is ignored in what follows, or, equivalently, assume the leisure decision is exogenous to the criminal activity decision, and that the time to be allocated between market work and criminal activity is the residual (total time in period-leisure).

We assume no capital markets exist so that individuals cannot borrow or lend money in any period. Total consumption in any given period then is purchased solely by contemporaneous income if the individual is incarcerated at no time during the period. This lack of the existence of capital markets is a limitation of the model; however, for purposes of

Table 1

	List of Symbols Used in Section 1.2
θ_t	Time devoted to criminal activity in period t $(0 \leq \theta_t \leq 1)$.
w_t	Legitimate work wage rate in period t.
H_t	The individual's criminal record as of time t (e.g., arrests, time in prison).
c^*	Consumption flow from incarceration.
$P(\theta_t)$	Probability of arrest in period t.
Y_{it}	Total monetarized returns from criminal activity in period t for individual i.
$F_i(Y_{it} \mid \theta)$	The conditional distribution function of criminal rewards.
$U(c)$	The utility of consumption level c.
δ_i	The parameter describing the conditional expectation of rewards in criminal activity for individual i $(E_i(Y \mid \theta) = \delta_i \theta)$.
$G(\delta)$	The distribution function of δ in the population.
τ	Sentence length if arrested.
β	Discount factor $(0 \leq \beta < 1)$.
η	The parameter describing the probability of incarceration function $(P(\theta) = \eta\theta)$.
V	The value of the individual's optimization problem in Section 1.2a.
α	The increment to wage rates for each period of non-incarceration.
$V(w_t)$	Value of the optimization problem in Section 1.2b for individual at wage level w_t.
S_t	Previous number of arrests as of period t.
$\tau(S_t)$	Sentence length function.
$V(S_t)$	Value of optimization problem in Section 1.2c for individual with arrest record S_t.

studying behavior in the criminally active subpopulation, it may not be entirely unrealistic.

Unlike legitimate activity, criminal behavior is "risky" in a particular sense. If an individual is caught engaging in criminal activity, he or she is incarcerated for at total of $\tau(H_{t-1})$ periods beginning with the period in which apprehension occurs, where H_{t-1} denotes the individual's criminal record through period t-1. Thus, if apprehension occurs in period t, the individual will be incarcerated for periods $t, t+1, t+\tau(H_{t-1}) - 1$. Note that sentence length is a deterministic function of the individual's previous criminal history, which at the beginning of period t is summarized by H_{t-1}. In general, it is reasonable to assume that the sentence length is an increasing function of the number of previous arrests, past time served in prison, or other observable characteristics of previous criminal activity. While incarcerated, the individual has a consumption level c^* each period.

The probability of being apprehended for criminal action in a period is a function of the amount of criminal activity engaged in over the period. This functional relationship is expressed as $P_t = P(\theta_t)$, where $P(\cdot)$ is monotonically increasing in θ_t and $P(0) = 0$, that is, if there is no criminal activity by the agent in the period, there is a zero probability of apprehension. It is not necessarily the case that $P(1) = 1$; that is, "full-time" criminals are not necessarily certain to be apprehended. In general $P(1) \leq 1$. Note that individual apprehension probabilities are only a function of current period activities, and not of criminal activities in previous periods.

To complete the specification of the choice set individuals face, we need to consider the potential rewards from criminal activity. Let the total monetary and psychic rewards from criminal activity in period t for individual i be denoted y_{it}. When the time allocation decision is made in period t, the final outcome or realization of y_{it} is unknown. Each individual does know the distribution of rewards he or she faces conditional on the time devoted to criminal activity. The conditional distribution function for individual i is given by $F_i(y_{it} | \theta)$. Unlike the other parameters of the problem, these conditional distribution functions differ across population members. This variation is meant to capture, in an admittedly limited way, the notion that individuals differ in their valuation of rewards from criminal activity. For all individuals, we assume that increases in θ, criminal activity, will increase the expected value of criminal rewards in the period. By the assumptions below, we do not need to consider the effect of θ on higher-order moments of the distribution.

Finally, we need to consider the total valuation of rewards from legitimate activities. Conditional on <u>not</u> being apprehended in period t, the expected utility of individual i in period t is given by

$$(1) \qquad E\, U_{it}(\theta_{it}, S) = \int U[(1 - \theta_{it})w + Y]\; dF_i(Y | \theta_{it}),$$

where it is assumed that $E_i(Y | \theta_{it})$ is bounded for θ_{it} in the unit interval, and where S (success) indicates that the individual was not apprehended in the period.

In what follows, we assume individuals are risk neutral, so that $U(x) = x$. This is done for reasons of tractability and because there seems no compelling reason to make differences in attitudes toward risk the basis of a model of differential criminal activity. Then (1) becomes

$$\text{(2)} \qquad E\, U_{it}(\theta_{it}, S) = (1 - \theta_{it})w_t + \int Y\, dF_i(Y \mid \theta_{it}).$$

The last term on the right-hand side of (2) is the expectation of criminal rewards in period t conditional on an activity level θ_{it}. We will consider the case in which this conditional expectation is linear, $E_i(Y \mid \theta_{it}) = \delta_i\, \theta_{it}$. This would be the case, for example, if the distribution of rewards was normal. The heterogeneity in individual valuations of criminal rewards is reflected in the fact that δ_i in the conditional expectation function varies across individuals in the population. The population distribution of δ is given by $G(\delta)$, defined over the interval $[\underline{\delta}, \overline{\delta}]$.

Now we can state the current period expected utility associated with a level of criminal activity θ_{it}. First note that the expected utility from action θ_{it} is given by $(1 - \theta_{it})w_t + \delta_i\, \theta_{it}$ and the probability of not being apprehended is $1 - P(\theta_{it})$. If the individual is apprehended and incarcerated, the utility yield is a certain c^*, and the probability of this occurring is $P(\theta_{it})$. Then expected utility in period t is

$$\text{(3)} \qquad E\, U_{it}(\theta_{it}) = \left[1 - P(\theta_{it})\right] \left[(1 - \theta_{it})w_t + \delta_i\, \theta_{it}\right] + P(\theta_{it})\, c^*.$$

Before proceeding to the three dynamic models, we note a few obvious restrictions on the parameters in this model. First of all, if $c^* > w_t$, there is no incentive not to engage in criminal behavior, for even if incarcerated, the individual would have a higher consumption value than when engaged in any level of market work. Second, assuming $c^* < w_t$, it must be the case that $\delta_i > w_t$ for at least some individuals in the population, or no criminal activity would be undertaken. These restrictions then are

(4a) $w_t > c^*$

(4b) $\bar{\delta} > w_t$.

Note that for any individual with a value of δ which satisfies the inequality $\delta \leq w_t$, no criminal activity will be undertaken in period t.[4] The analyses below only pertain to individuals with $\delta > w_t$; all others will optimally choose not to engage in criminal activity. We now turn to the consideration of dynamic behavior under three different specifications of constraints on criminal choices.

2.1 The Constant Wage Case

To begin we consider the case in which the wage of each individual in the population is fixed over time: $w_t = w$, $t = 0, 1, \ldots$. We will also begin by assuming that conditional on apprehension, sentence length is the same for all individuals, regardless of previous criminal history, so $\tau(H_{t-1}) = \tau$, $t = 1, 2, \ldots$. Since we assume individuals are infinitely lived, and by these assumptions the choices individuals face are constant over time (but may differ across individuals), each individual will denote the same amount of time to criminal activity each period in which he is not initially incarcerated. For an individual, his constant rate of criminal activity θ^* will be a function of the parameters characterizing his preferences and constraints. In this first simple model then, $\theta^* = \theta^*(c^*, \delta, P(\cdot), w, \tau)$. (The individual subscript i has been dropped for notational simplicity). We now turn to an investigation of the function θ^*.

Denote the value of being free (not incarcerated) at the beginning of any period by V. Conditional on his choice of θ in the period, his expected utility given that he is not incarcerated is $(1 - \theta)w + \delta\theta + \beta V$. The term βV is interpreted as follows. If the individual is not incarcerated in this period, he will be free to make a time allocation decision next period. By the structure of this problem, the value of the decision is given by V. But rewards in the future are not perceived by individuals to be as valuable as rewards today. The rate at which individuals discount future rewards is given by the discount factor β, $0 \leq \beta < 1$. (If $\beta = 0$, individuals completely ignore the effect of their current actions on future choices. As β approaches 1, individuals consider current and future rewards as virtually perfect substitutes.) Thus the value of being free next period evaluated as of this period is βV. The probability of not becoming incarcerated is $1 - P(\theta)$.

The "value" of becoming incarcerated during the period is determined in the following way. If incarcerated, the individual will serve τ periods in prison, beginning today. The value of being in prison in the current period is c^*; as of today, the value of being in jail next period is βc^*; and for m periods from now is $\beta^m c^*$. Then the utility yield during the period of incarceration is $\sum_{i=0}^{\tau-1} \beta^i c^*$. In addition, the individual will be free to allocate his time optimally in τ periods--the value of this is $\beta^\tau V$. Then the total value of incarceration is $\sum_{i=0}^{\tau-1} \beta^i c^* + \beta^\tau V$. The probability of incarceration is $P(\theta)$.

When we combine all the elements discussed above, the maximum value of the individual's time allocation problem in all periods when he is not incarcerated as of the beginning of the period is given by

$$(5) \qquad V = \max_{0 \leq \theta \leq 1} \left\{ [1 - P(\theta)][(1 - \theta)w + \delta\theta + \beta V] + P(\theta)[\sum_{i=0}^{\tau-1} \beta^i c^* + \beta^\tau V] \right\}.$$

To simplify discussion, we make a further functional form assumption. Let the conditional probability of apprehension, $P(\theta)$, be given by $P(\theta) = \eta\theta$, $0<\eta\leq 1$. Then we have

$$(5') \qquad V = \max_{0 \leq \theta \leq 1} \left\{ [1 - \eta\theta][(1 - \theta)w + \delta\theta + \beta V] + \eta\theta[c^* \sum_{i=0}^{\tau-1} \beta^i + \beta^\tau V] \right\}.$$

Denote by $\tilde{\theta}^*$ the amount of time devoted to criminal activity not taking into account the restriction that this is a proportion lying in the unit interval. Then $\tilde{\theta}^*$ is given by

$$(6) \qquad \tilde{\theta}^* = [2\eta(\delta - w)]^{-1} [\delta - w(1 + \eta) - \eta\beta V + \eta c^* \sum_{i=0}^{\tau-1} \beta^i + \eta\beta^\tau V].$$

The solution to (5') is denoted θ^*. Then

$$(7) \qquad \theta^* = \begin{cases} 0 & \text{if } \tilde{\theta}^* \leq 0 \\ \tilde{\theta}^* & \text{if } 0<\tilde{\theta}^*<1 \\ 1 & \text{if } \tilde{\theta}^* \geq 1. \end{cases}$$

If $\theta^* = 0$ or $\theta^* = 1$, we will say the individual's time allocation problem yields a corner solution. If $\theta^* = 0$, he is always engaging in legitimate activity; if $\theta^* = 1$, he is a "full-time" criminal. An interior solution exists if $0<\theta^*<1$; in this case the individual devotes some time to criminal activity and some time to legitimate activity.

For this model it is possible to find a closed-form solution in the following manner. Note that V is defined by

$$(8) \qquad V = [1 - \eta\theta^*][[1 - \theta^*]w + \delta\theta^* + \beta V] + \eta\theta^*[c^* \sum_{i=0}^{\tau-1} \beta^i + \beta^\tau V].$$

Solving for V we obtain

$$(9) \qquad V = \left[1 - \beta(1 - \eta\theta^*) - \eta\theta^*\beta^\tau\right]^{-1} \left[(1 - \eta\theta^*)\left((1 - \theta^*)w + \delta\theta^*\right) + \eta\theta^* c^* \sum_{i=0}^{\tau-1} \beta^i\right].$$

This can be written as

$$(9') \qquad V = \frac{a_0 + a_1\theta^* + a_2(\theta^*)^2}{b_0 + b_1\theta^*},$$

where $a_0 = w$, $a_1 = \delta - w(1 + \eta) + \eta c^* \sum_{i=0}^{\tau-1} \beta^i$, $a_2 = \eta(w - \delta)$, $b_0 = 1 - \beta$, and $b_1 = \eta(\beta - \beta^\tau)$. From equation (6), write

$$(6') \qquad \theta^* = c + dV,$$

where $c = [2\eta(\delta - w)]^{-1} [\delta - w(1 + \eta) + \eta c \sum_{i=0}^{\tau-1} \beta^i]$ and $d = [2(\delta - w)]^{-1} [\beta^\tau - \beta]$. Substituting (9') into (6'),

$$(\theta^*)^2 + e\theta^* + q = 0,$$

where $e = 2b_0/b_1$ and $q = [b_1 a_2]^{-1} [a_1 b_0 - b_1 a_0]$. Thus the solution for θ^* is given by

$$(10) \quad \theta^* = - b_0/b_1 + [(b_0^2 a_2 - a_1 b_0 b_1 + b_1^2 a_0)/b_1^2 a_2]^{1/2}.$$

Since a closed-form solution is available for the proportion of time spent in criminal activity, it is straightforward to determine the qualitative effect of changes in the parameters (η, β, w, δ, c^*, τ) on behavior. Qualitatively, the following results hold:

$$(11) \quad \frac{\partial\tilde{\theta}^*}{\partial\delta} \geq 0; \ \frac{\partial\tilde{\theta}^*}{\partial c^*} \geq 0; \ \frac{\partial\tilde{\theta}^*}{\partial\tau} \leq 0;$$

$$\frac{\partial\tilde{\theta}}{\partial\eta} \leq 0; \ \frac{\partial\tilde{\theta}^*}{\partial w} \leq 0.$$

Verbally, an increase in the expected marginal rate of return to criminal activity, δ, results in an increase in the rate of criminal activity. The rate of criminal activity also is increasing in the utility associated with "failure" (incarceration), which is given by parameter c^*. As punishments increase in length (τ), criminal activity declines. Increases in the marginal arrest rate (η) result in decreases in crime rates. [An increase in the direct opportunity cost of crime, the wage rate in legitimate work (w), causes decreases in the rate of crime.]

Results of this type have been obtained previously in a number of static rational choice models of criminal behavior. In fact, if sentence length τ is equal to 1, this model reduces to a series of static optimization problems. By allowing $\tau \geq 2$, individuals face one of two different choice problems at the beginning of each period. If they are not incarcerated at the beginning of the period, they choose the amount of time to engage in criminal activity, θ, and as noted above, in this model, they will always set θ to the same value. If they begin the period incarcerated, their current period utility is predetermined at the value c^*.

The only parameter that reflects the dynamics of the problem aside from the sentence length τ is the discount factor β. In this model, the sign of $\partial\theta^*/\partial\beta$ is ambiguous. This partial derivative can be computed in a straightforward manner, but the result is not particularly enlightening. The intuition is basically the following. In any period in which the individual is free initially, his expected current period utility is given by equation (3), and by the assumptions of the model, $E\ U_i(\theta_i^*) > c^*$ for each individual i in the population. If the sentence length is τ periods, the difference in expected utility of freedom versus incarceration is $(1 + \beta + \beta^2 + \ldots + \beta^{\tau-1})(E\ U_i(\theta_i^*) - c^*)$. Holding constant θ_i^*, an increase in β increases this cost. However, for any finite length sentence, τ, an infinitely lived individual (or an individual with a sufficiently long but finite life) will eventually be released. The value of being free at the time of the release is $\beta^\tau V$.

As β approaches 1, $\beta^{\tau} \to \beta$ so the value of being free at the beginning of the period τ periods from the present ($\beta^{\tau}V$) approaches the value of being free next period (βV). At the same time, as $\beta \to 1$, the value of the optimization problem goes to infinity. Thus the penalty $(1 + \beta + \ldots + \beta^{\tau-1})(E\ U_i(\theta_i^*) - c^*)$ becomes insignificant, and this results in increases in criminal activity. Which effect will dominate depends on all the parameter values in the model.

By the assumptions of this model, individuals commit a constant rate of crimes over their lifetimes, which is contradictory to the empirical evidence that exists. In the next section we produce a model in which criminal activity decreases on average as individuals age.

2.2 Accumulation of Human Capital in Legitimate Activity

In the constant wage case we obtained the result that the proportion of time nonincarcerated individuals devoted to criminal activity remained constant as they aged. There are several ways in which the simple model presented in Section 2.1 could be modified so as to produce the result that the crime rate is a decreasing function of age. One obvious modification is to allow the returns from legal and/or illegal activity to be age dependent. Intuitively if the difference between returns from legitimate work and expected returns from criminal activity diminish over time, other things equal, we may observe a decrease in the crime rate with age. (Recall that it was necessary to assume that the expected returns from crime were strictly greater than the legitimate wage if we were to observe any criminal activity. As the legitimate wage approaches the expected returns from crime, we will observe a continuous decline in the crime rate of an individual.)

The approach we take in this section is to hold the expected returns from criminal activity constant but allow the legitimate wage to change systematically over the life cycle as a result of individual behavior and random events. While it would be desirable to allow the expected returns from criminal activity to also vary systematically over the life cycle, such an extension would add greatly to the complexity of the model. Furthermore, what is really of interest is the difference between expected rewards from criminal activity and legitimate work. Thus it is somewhat inconsequential whether we model the change in this difference as resulting from shifts in the legitimate wage, the expected returns from crime, or both.

There exists a voluminous literature in the area of human capital accumulation. For a statement of the general theory, see Becker (1975). We assume here that there is no accumulation of crime-specific human capital--that is, individuals do not become more proficient criminals as they acquire criminal experience. Market wage rates do increase as individuals acquire market experience however. We will characterize this dependency in the following way. In period t, when the amount of criminal activity is given by θ_t, we will say the individual accumulates a total of $(1 - \theta_t)$ units of experience if he is not incarcerated during the period. If he is incarcerated in the period he acquires no market experience. Similarly, if he is incarcerated at the beginning of the period--he is not free--he accumulates no market experience. The amount of market human capital the individual has at the beginning of period t will be denoted by h_t. Then the wage rate an individual faces in the period will be assumed equal to $h_t (w_t = h_t)$. The amount of human capital the individual possesses at the beginning of period t is defined in the following way. First define a variable

$$\hat{\theta}_k = \theta_k^* \text{ if the individual was not incarcerated during period k}$$

$$= 1 \text{ if he was incarcerated during period k.}$$

Then the total amount of market experience the individual has at the beginning of period t is

$$\sum_{k=1}^{t-1} (1 - \hat{\theta}_k).$$

Market human capital is assumed to be a simple transformation of market experience,

$$(12) \qquad h_t = g\left(\sum_{k=1}^{t-1} (1 - \hat{\theta}_k)\right),$$

where g is a monotonically increasing function; human capital is increasing in labor market experience.

Now the choices an individual can make at any time t depend on his past allocation of time, $\{\theta_k^*\}_{k=1}^{t-1}$, in all periods when he was free, and luck--that is, how often he was incarcerated in the past. These are the sources of variation in the sequence $\{\hat{\theta}_k\}_{k=1}^{t-1}$ which determines beginning of period t human capital, and hence the period t wage rate.

At any age, $t = 1, 2, \ldots$, individuals will in general be differentiated according to their stock of human capital. Consider an individual making a time allocation decision in period t. His choice of rate of criminal activity will depend on his current period wage rate, h_t. This wage rate changes over time and is called a state variable. An individual in state h_t faces the optimization problem

$$(13) \quad V(h_t) = V(w_t) = \max_{0 \leq \theta_t \leq 1} \Big\{ [1 - \eta\theta_t][(1 - \theta_t)w_t + \delta\theta_t + \beta V(w_{t+1}(\theta_t, w_t))] + \eta\theta_t[c^* \sum_{i=0}^{\tau-1} \beta^i + \beta^\tau V(w_t)] \Big\}.$$

where $w_{t+1}(\theta_t, w_t)$ denotes the fact that given the wage rate in period t, w_t, the time allocated to criminal activity, θ_t, and the fact that the individual was not incarcerated during the period, the period t+1 wage is known with certainty. The function $w_{t+1}(\theta_t, w_t)$ is decreasing in the amount of time spent in criminal activity and increasing in the previous wage rate. Note that if an individual is incarcerated in period t, when he is released in period $t+\tau$ he will be able to work at the same wage as in period t. Thus we have assumed an absence of stigma--the effect of jail on wages is simply an absence of growth, not a decline.

In the changing wage case there exist a number of additional costs of criminal activity. To review the structure of the model, we will list them all:

1) In the current period t, if the individual is not incarcerated, the opportunity cost of crime is simply foregone market work, which is remunerated at rate w_t.

2) In period t, increased criminal activity also increases the probability of incarceration. Minus the difference between the level of expected utility as a free individual and that obtained as a prisoner multiplied by the increase in the probability of being incarcerated is an additional cost of increased criminal activity.

3) Conditional on the current wage rate w_t, increases in criminal activity decrease next period's wage w_{t+1} (given no incarceration in period t) owing to foregone human capital accumulation. Since the expected utility of all individuals is an increasing function of the market wage in all periods, the lower future wage rate must lower expected utility levels in future periods.

4) Increased criminal activity leads to an increased probability of incarceration, and while incarcerated for τ periods the individual does not accumulate any market capital. This represents a permanent wage reduction in future periods, or a persistent effect of incarceration.

Solving (13) turns out to be quite difficult in practice, even for simple forms of the human capital accumulation function g. Therefore for the remainder of this section we confine our discussion to the following special case. We will assume that as long as an individual is not incarcerated during period t, his wage will increase by α in period t+1. Then

$$w_{t+1} = w_t + \alpha \qquad \text{(given no incarceration in period t).}$$

Given that the individual is not incarcerated in period t, his period t+1 wage is independent of θ_t $(w_{t+1}(\theta_t, w_t) = w_{t+1}(w_t))$. The cost referred to in point (3) above is absent. However, point (4) is still operative--increased crime increases the probability of incarceration that is associted with foregone human capital accumulation.

With this simplification, (13) can be rewritten

$$(13') \quad V(w_t) = \max_{0 \leq \theta_t \leq 1} \{[1 - \eta\theta_t]](1 - \theta_t)w_t + \delta\theta_t + \beta V(w_t + \alpha)]$$

$$+ \eta\theta_t[c^* \sum_{i=0}^{\tau-1} \beta^i + \beta^\tau V(w_t)]\}.$$

Now the individual's time allocation problem depends on the set of parameters in the constant wage case plus the wage growth parameter α. Unlike the constant wage case, it is not possible to find closed-form solutions for $\tilde{\theta}^*(w_t)$ or $V(w_t)$, so numerical methods must be used to investigate quantitative properties of these functions. However, all the comparative statics results in (11) hold for the changing wage case, and in addition, $\partial\tilde{\theta}_t^*/\partial\alpha < 0$--the larger the wage increment, the lower the crime rate, for the larger is the opportunity cost of incarceration.

Finally, to illustrate the individual-level and aggregative characteristics of this model, we computed some numerical examples. These results are not meant to constitute any exhaustive study of the function (13'), but merely serve to demonstrate the types of criminal careers that can be generated by this simple model. We should say that the parameter values selected for this illustration were not chosen after an exhaustive search. It appears that this model can generate "interesting" career patterns (i.e., not all corner solutions) without extensive search over the parameter space.

The actual parameter values chosen were arbitrary. The initial wage level, w_1, is set to .5. Then we impose condition (4a) by setting $w_1 > c^*$, and in particular set $c^* = 0$. The wage increment, α, is set to .05. The discount factor, β, is equal to .8, the arrest parameter, η, is set to .5, and the sentence length, τ, is set to 3 periods. All individuals face these same parameters, however we assume that two distinct values of δ exist in the population. The conditional expectation parameter assumes the value 3 for 50 percent of the population and assumes the value 2 for the other 50 percent. We will refer to the $\delta = 3$ individuals as "high crime types" and the $\delta = 2$ individuals as "low crime types."

In Table 2 we present the amount of time devoted to criminal activity as a function of the beginning-of-period wage level for both population groups. Note that both types devote substantial amounts of time to crime at initial wage level .5. Criminal activity quickly drops off for the low crime types--no criminal activity occurs at a wage of .8. This is not the case for the high crime types--criminal activity only ceases at a wage of 1.25. In particular, at wage .8, when low crime types cease criminal activity, the high crime types still devote 43 percent of their time to criminal activity.

Using the decision rule given in Table 2, we can investigate patterns of individual offending in the population. The procedure used is straightforward. Consider the case of an individual who is a low crime type. In period 1, his wage is .5 and consequently he spends .43 percent of his time in criminal activity. Since $\eta = .5$ and the probability of incarceration is $\eta\theta$, he is arrested in the first period with probability .215. A random number generator is used to determine the outcome of this chance event. If arrested, he is sent to jail in period 1 and not released until period $\tau+1$(period 4 in the case $\tau = 3$). If not arrested, he is free at the beginning of period 2 with a wage of .55. The process

Table 2

Time Allocation to Criminal Activity,
Wage Growth with Constant Sentence Length

Wage Level	Low Crime Types	High Crime Types
.5	.4305	.6007
.55	.3726	.5774
.6	.3082	.5524
.65	.2370	.5258
.7	.1590	.4972
.75	.0747	.4663
.8	0.0	.4330
.85		.3969
.9		.3579
.95		.3156
1.0		.2698
1.05		.2205
1.1		.1675
1.15		.1111
1.2		.0515
1.25		0.0

is repeated in this manner for 50 periods of life for each of 1000 "individuals" comprising the low and high crime type groups.

In Table 3 we display the total amount of crime committed by the cohort, the total number of arrests, and the beginning-of-period jail population. Note that initially individuals who are high crime types are responsible for a bit less than 60 percent of total crime. By period 10, they are responsible for 90 percent of total crime, and by period 20, they are responsible for virtually all crime. This obviously has implications for identification of high and low type offenders. Classification of individuals arrested in period 1 into low and high crime type groups involves a substantial amount of error. An individual arrested in period 20 however, may be classified with virtual certainty as a high crime type. An even more accurate classification can be made if the wage rate of arrested individual is available. From Table 2 we know that if an individual with a wage greater than or equal to .8 is arrested, he must be a high crime type. At wages less than .8 the relative likelihoods are given by the ratio of column 3 to column 2.

Note that while these results indicate the potential for identification of population subgroup members, no individual is incorrigible. By altering the wage rates of high crime types, or lowering their expected return from criminal activity, these individuals, once identified, can be induced to spend the same or less time in crime than the other group in the population.

Table 3

Aggregate Crime Statistics in Simulated Population,
Wage Growth with Constant Sentence Length
(1000 Individuals in Each of
High and Low Crime Type Groups)

Period	Total Crime		Arrests		Jail Population	
	Low	High	Low	High	Low	High
1	430.4	600.7	214	301	0	0
2	292.9	403.6	143	201	214	301
3	198.2	275.1	92	144	357	502
4	222.7	366.9	112	188	235	345
5	192.6	365.8	92	200	204	332
6	138.6	322.2	67	161	204	388
7	105.7	328.6	54	158	159	361
8	92.1	343.2	48	151	121	319
9	64.3	334.5	30	172	102	309
10	40.3	314.9	22	158	78	323
11	35.4	300.8	17	150	52	330
12	24.6	298.3	14	164	39	308
13	13.5	278.6	8	115	31	314
14	11.4	277.5	11	153	22	279
15	7.4	264.5	1	126	19	268
16	4.0	235.6	3	100	12	279
17	4.4	237.6	0	114	4	226
18	2.6	217.8	1	112	3	214
19	1.4	186.3	2	92	1	226
20	.6	172.3	1	86	3	204
25	0.0	89.2	0	48	0	121
30	0.0	42.4	0	20	0	47
35	0.0	10.2	0	7	0	17
40	0.0	1.5	0	0	0	8
45	0.0	.4	0	1	0	1
50	0.0	.2	0	0	0	0

2.3 Increasing Penalties for Criminal Activity

In the last section we demonstrated that with the benefits of legitimate market work increasing on average over the life cycle, the rate of criminal activity would decrease with age. In this section we briefly discuss the manner in which differential sentencing would produce the same relationship between age and the rate of criminal activity.

As in Section 2.1, we assume that legitimate market wages w are constant over time so that we may isolate the sentencing effect. Previously we asssumed that sentence lengths, τ, were constant, which is obviously not the case in practice. Not only do sentence lengths differ by type of crime, the length of a sentence typically depends on the number of times the individual has previously been convicted of criminal activity. We will continue to confine our attention to the one crime type case. We will only be concerned with modelling the dependence of sentence length on the number of past convictions for this one type of crime.

In this model we define a new state variable, S_t, which denotes the number of previous convictions as of the beginning of period t. Sentence length is no longer a constant, but is a function $\tau(S_t)$, where it is reasonable to assume $\tau(0) \leq \tau(1) \leq \ldots$. Now the rewards for legitimate work are the same in all periods, as are the rewards from criminal activity <u>if successful</u>. Only the punishments change as a consequence of changes in the state variable S_t. The individual's time allocation problem is

$$(14) \quad V(S_t) = \max_{0 \leq \theta_t \leq 1} \{ [1 - \eta\theta_t][(1 - \theta_t)w + \delta\theta_t + \beta V]$$

$$+ \eta\theta_t [c^* \sum_{i=0}^{\tau(S_t)-1} \beta^i + \beta^{\tau(S_t)} V(S_t + 1)] \}.$$

Corresponding to this problem there exists a solution $\theta^*(S_t)$. The ordering of the solutions is $\theta^*(0) \geq \theta^*(1) \geq \theta^*(2) \geq \ldots$. The larger are the differences in the sentence length function $\tau(k) - \tau(k-1)$, $k = 1, 2, \ldots$, the larger are the differences $\theta^*(k) - \theta^*(k-1)$. We note parenthetically that large changes in τ as a function of sentence length may result in individuals who originally devote a substantial amount of time to criminal activity, eventually switching out of crime completely.

To illustrate the characteristics of criminal careers generated by this model an example similar to the one in Section 1.2b is provided. All parameter values are exactly the same with the exception of the sentence length τ. In this model, τ is set to 1 if the individual has <u>no</u> prior convictions. It is set to 5 if the individual has <u>any</u> prior convictions.

The decision rules are presented in Table 4. The amount of criminal activity for low and high crime types is greater than was the case in Table 2 <u>conditional</u> on no previous arrests. This increased activity results in increased arrest probabilities, however, and after one arrest individuals devote less time to criminal activity than was the case in Table 2. After one arrest, an individual of the low crime type receiving a wage of .5 will spend only about one-third as much time in criminal activity as was previously the case. High crime types also substantially

Table 4

Time Allocation to Criminal Activity,
Wage Growth with Varying Sentence Length

Wage Level	Low Crime Types		High Crime Types	
	No Arrests	Some Arrests	No Arrests	Some Arrests
.5	.5711	.2176	.6693	.4786
.55	.5525	.1326	.6513	.4466
.6	.5374	.0406	.6330	.4117
.65	.5242	0.0	.6145	.3734
.7	.5098		.5962	.3313
.75	.4937		.5786	.2851
.8	.4755		.5621	.2345
.85	.4544		.5472	.1796
.9	.4297		.5343	.1204
.95	.4004		.5233	.0574
1.0	.3652		.5137	0.0
1.05	.3222		.5042	
1.1	.2697		.4941	
1.15	.2058		.4832	
1.2	.1291		.4714	
1.25	.0392		.4587	
1.3	0.0		.4449	
1.35			.4298	
1.4			.4131	
1.45			.3945	
1.5			.3736	
1.55			.3499	
1.6			.3228	
1.65			.2915	
1.7			.2553	
1.75			.2132	
1.8			.1646	
1.85			.1089	
1.9			.0463	
1.95			0.0	

reduce criminal activity after one arrest, but not to the same degree as low crime types.

Aggregate statistics are presented in Table 5. Compared to Table 3, we see that a greater amount of crime is forthcoming in the first few periods in the varying-sentence-length model, but eventually total crime is reduced as more individuals are subject to the stiffer sentence $\tau = 5$. The jail population is substantially smaller over the life of the cohort in the varying-sentence-length model.

2.4 Identification of Structural Models

The models proposed in Sections 2.1 - 2.3 were primarily designed to illustrate how various empirical regularities, such as the decline of crime rates with age, may be generated from dynamic behavioral models. As discussed in Section 1, such structural models may be preferable to less behaviorally motivated statistical models in that all parameters have relatively clear interpretations. Structural models are probably not useful when their structure precludes them, a priori, from being able to reproduce salient empirical regularities.

If these structural models are to prove useful empirically, we must of course be able to obtain consistent estimates of all or most of the parameters in the decision rules. The first consideration is one of identification. What types of data are required to estimate one of these models? Let us consider the model presented in Section 2.3 for example.

The model with increasing sentence lengths is described by the following set of parameters: η, w, δ, β, c^*, $\tau(\cdot)$. Identification may proceed in the following way. First recognize that the consumption value

Table 5

Aggregate Crime Statistics in Simulated Population,
Wage Growth with Varying Sentence Length,
(1000 Individuals in Each of High
and Low Crime Type Groups)

Period	Total Crime		Arrests		Jail Population	
	Low	High	Low	High	Low	High
1	571.1	669.3	265	368	0	0
2	463.7	587.8	213	292	0	0
3	351.3	488.9	190	264	27	81
4	228.2	386.8	103	195	61	193
5	152.4	302.3	72	158	70	296
6	112.1	229.3	55	122	70	399
7	87.7	209.4	43	91	43	410
8	67.7	207.8	38	103	12	364
9	45.1	189.6	24	111	8	346
10	31.2	164.6	22	74	11	337
11	20.5	139.2	12	79	11	311
12	14.8	105.4	8	52	8	315
13	10.8	103.9	4	45	3	278
14	7.0	102.4	3	52	0	226
15	3.7	87.4	1	44	0	210
16	1.1	72.3	0	30	0	183
17	0.0	54.6	0	23	0	164
18	0.0	52.0	0	35	0	145
19	0.0	48.1	0	31	0	128
20	0.0	40.9	0	18	0	116
25	0.0	19.2	0	10	0	61
30	0.0	9.2	0	3	0	30
35	0.0	4.1	0	2	0	9
40	0.0	1.2	0	1	0	2
45	0.0	.4	0	1	0	0
50	0.0	.1	0	0	0	0

of being in prison, c^*, is arbitrary. Setting it to a given value essentially fixes the location of the utility index. It seems most natural to set $c^* = 0$. Now the rate of arrest, conviction, and incarceration may be computed from victimization surveys, which give an estimate of the total number of crimes committed (of a particular type). This combined with the number of individuals incarcerated for the crime will yield an estimate of η. The sentencing function $\tau(\cdot)$ is also relatively straightforward to compute. Either actual sentencing records may be used, or official guidelines, when available.

The parameters β, w, and δ present more of a challenge. Often the value β is not estimated in analyses of this sort--it is merely set to a "reasonable" value, typically .9 or .95. Since criminally active individuals are often thought to discount the future rather heavily (that is, have low values of β), in an analysis like the present one it may be of interest to estimate β. This parameter is in principle identified in this model, at least if we are able to observe θ_t^*--the proportion of criminal activity in period t. The other individual-level data we will need for purposes of identification are wage rates. Wage rates are obviously not identical over time and individuals--neither are they identical over time for the same individual. We can incorporate this observation by assuming $w_{it} \sim N(\underset{\sim}{X}'_{it}\, \underset{\sim}{\gamma}, \sigma^2)$, for example, where $\underset{\sim}{X}'_{it}$ is a vector of individual characteristics at time t and $\underset{\sim}{\gamma}$ and σ^2 may be estimated. In periods when individuals are full-time criminals no wage will be observed--but by making the distribution assumption on w_{it}, data from such periods will still be informative for $\underset{\sim}{\gamma}$, σ^2.

Estimation of the parameter δ, or its distribution in the population, is the most difficult. It is not necessary to measure the returns from criminal activity in order to estimate this parameter. One could proceed in the following fashion. First assume a form for the population distribution of δ, say $M(\delta, \underset{\sim}{\xi})$ where $\underset{\sim}{\xi}$ is a parameter vector which characterizes M. The likelihood of observing w_{it} and θ^*_{it} in a period can be constructed conditional on a value of δ. By taking the expected value of this conditional likelihood with respect to the distribution to δ, we can form an unconditional likelihood which depends on the parameters $(\underset{\sim}{\gamma}, \sigma^2, \beta, \underset{\sim}{\xi})$. My conjecture is that for identification of $\underset{\sim}{\xi}$, β must be fixed. But note that in this analysis it is possible to estimate a rather abstract but interesting distribution $M(\delta, \underset{\sim}{\xi})$, even if we assume criminal rewards are not measurable, or even operationally definable.

3. ECONOMETRIC MODELS OF CRIMINAL CAREERS

In the previous section we developed dynamic models of the criminal career based on optimizing behavior. As discussed in Section 1.1, there are advantages and disadvantages to the estimation of such highly structured models. In short, the principal advantage is unambiguous interpretation of parameter estimates and statistical tests. The principal disadvantages are the complicated computational algorithms required for estimation and the typically poor "explanatory" power of such models. Given our current level of understanding of the simple statistical properties of the criminal career process, perhaps it is beneficial to work with econometric models less closely linked with a specific behavioral model, but that allow for statistical associations precluded in any tractable decision-theoretic model. Actually, the choice is not between one approach or the other. Both can and should be used in any systematic study of the criminal career.

In this section we briefly outline the relevant theory and present a relatively general framework in which parameters of continuous time behavioral models may be estimated. We will be concerned with the econometric and statistical properties of continuous time models for the most part.

To fix ideas, consider a continuous time, discrete state space stochastic process X, where the state space consists of the nonnegative integers, $S = Z^+$, and where the parameter set $T = [0, \infty)$. The state of the process at time t, s_t, indicates the number of times some event has occurred from the origin of the process, normalized at 0 without loss of generality, through time t. For example, say life began at time 0, and there is only one type of crime individuals can commit. Then s_t indicates the number of times an individual has committed this crime as of age t. Let the times at which crimes occur be given by $\tau_1, \tau_2, \ldots$. Then the number of crimes previously committed as of time t (s_t) is equal to s^* if and only if $\tau_{s^*} \leq t < \tau_{(s^*+1)}$. Also, define the duration until the first crime as $w_1 = \tau_1 - 0 = \tau_1$. The duration of the kth spell, i.e., the elapsed time between crimes k and k-1, is $w_k = \tau_k - \tau_{k-1}$.

We can characterize the stochastic process in a number of alternative ways. For the most part in this section we will discuss the interval specification of the process. Then, starting from time 0, we will characterize the process through the joint distribution of intervals between events, $F_n(w_1, w_2, \ldots, w_n)$, for $n = 1, 2, \ldots$.

In terms of a specific application to criminal career analysis, the formalism above reduces to the following. Say an individual is born at time 0 and lives to age $T(T \leq \infty)$. At each instant of life t, $0 \leq t \leq T$, the individual either commits a crime or doesn't. The length of time between successive criminal acts is in general not constant. Loosely speaking, the agent's propensity to commit a crime at some particular time $\tilde{t}$ depends not only on his or her "normal" rate of crime commission and the elapsed time since the last crime was committed, but also on the opportunities for crime commission that exist at that particular moment. Thus even if an individual would "normally" be highly likely to commit a crime at $\tilde{t}$, the fact that a policeman happened to be in close proximity would probably induce a postponement to a later date. Or the fact that an attractive mark appears may induce a crime before we would normally expect one. Unanticipated or anticipated changes in the choice sets of individuals will cause variations in criminal behavior over the life cycle, as the preceding section demonstrated. To the extent that these changes are not anticipated by the individual or observable to the analyst, the process of crime commission must be considered to be random. Then the length of time from the beginning of life to the time of commission of the first crime is w_1, which is a random variable. The distribution of w_1 is given by $F_1(w_1)$. Analagously, the length of time between the first and second commission, w_2, has distribution $F_2(w_2)$, and in general the duration of time between crime $(i-1)$ and i is $F_i(w_i)$. The joint distribution over the first n crimes is given by $F_n(w_1, w_2, \ldots, w_n)$, as stated above.

To be useful for purposes of statistical (or theoretical) analysis of the criminal career, some structure must be imposed on the general joint distribution $F_n(w_1, w_2, \ldots, w_n)$. A natural starting point is to assume that, for a given individual, all spell lengths are independently distributed, i.e., the joint distribution of durations $w_1, w_2, \ldots, w_n$ can be written as

$$F_n(w_1, w_2, \ldots, w_n) = F_1(w_1)F_2(w_2) \ldots F_n(w_n),$$

$$\text{for } n = 1, 2, \ldots.$$

Simply stated for the case n = 2, this implies the length of interval 1 does not alter our assessment of the likelihood of observing any particular value of second spell duration.

By adding another assumption concerning the joint distribution of the spell lengths, we will produce a class of models often used in engineering and increasingly in the social sciences. If we assume that the distribution functions have the same (identical) form,

$$F_1(s) = F_2(s) = \ldots F_n(s); \qquad s \geq 0,\ n = 1, 2, \ldots,$$

then we can write the joint distribution of the first n spells as

$$F_n(w_1, w_2, \ldots, w_n) = \prod_{i=1}^{n} F(w_i).$$

A point process in which the spell lengths are independently and identically distributed (i.i.d.) is a renewal process.

If the common (to all spells) duration distribution is everywhere differentiable (as we will assume throughout this paper), there exists an associated probability density function f(w). A renewal process can be completely characterized by F(w) or f(w), if it exists. Alternatively, it can be characterized by its hazard rate function h(w), which is defined as

$$h(w) = \frac{f(w)}{1 - F(w)}, \qquad w \geq 0.$$

The hazard rate function is the conditional density of duration times given the individual has not committed a crime for a period of length w. The hazard rate function h will be utilized in the econometric model formulated below.

One of the most important characteristics of a duration density from both a behavioral and statistical perspective is the degree and type of duration dependence exhibited. Duration dependence is most easily investigated through the hazard rate function. Simply differentiate h(w) with respect to w, dh/dw. If

$$\left.\frac{dh(w)}{dw}\right|_{s} \gtreqless 0,$$

we say that the hazard rate function (or density) exhibits positive, no, or negative duration dependence when evaluated at duration s. If the sign of the derivative is the same for all $s \in (0, \infty)$, we say that the hazard or density exhibits monotonic duration dependence. If the signs switch at least once, duration dependence is nonmonotonic. The only duration density that exhibits no duration dependence over the entire interval $(0, \infty)$ is the exponential, $f(w) = \phi \exp(-\phi w)$, $\phi > 0$.

Parameterizing the hazard directly has many econometric advantages that are discussed below. One of these is that information from incomplete spells, those which began during the sample period but had not been completed when the sample period ended, may be incorporated into the estimation procedure in a straightforward way. In actuality, individuals are only observed over some portion of their lifetime. Let the sampling period be the interval $(0, \ell]$, and assume over this interval the individual is observed to commit m crimes. For the pure renewal process described above, we know that the mth event occurred at time τ_m, however we did not observe the time at which the $(m+1)^{st}$ event occurred. We do know that this event had not occurred by the end of the sample, ℓ. This occurs with probability $1 - F(\ell - \tau_m)$. It is easy to show that this quantity, referred to as the survivor function, is equal to $\exp\{-\int_0^s h(u)du\}$, where $s = \ell - \tau_m$.

Environments are, of course, highly nonstationary, and at a single point in time there exist substantial amounts of heterogeneity in the population with respect to budget sets and preferences. Renewal processes can still provide a useful framework with which to analyze dynamic economic behavior if we generalize them so as to incorporate some forms of nonstationarity and heterogeneity. We may retain the i.i.d. assumptions regarding the density of duration times, but make the parameters describing the duration density functions of observable and unobservable individual characteristics. These characteristics may change over time. For example, we may write the conditional hazard function as

$h(w_{ik} \mid \underset{\sim}{Z}_i(\tau_{ik} + w_{ik}); \underset{\sim}{\theta})$, where k indexes the serial order of the spell, w_{ik} is the duration of the kth spell for individual i, $\underset{\sim}{Z}_i(\cdot)$ is an individual-specific vector of observable and unobservable sources of heterogeneity which can be time-varying, and $\underset{\sim}{\theta}$ is a conformable parameter vector.

In the case of the pure unconditional renewal process first described, the density of duration times f(t) could be estimated by parameteric or nonparametric methods simply from a sufficiently large number of completed spells for <u>one</u> individual. Once we allow for conditioning on a set of individual characteristics $\underset{\sim}{Z}_i$, some of which may be time invariant, it is clear that in order for all elements of $\underset{\sim}{\theta}$ to be estimable we will need observations for many individuals. The econometric model we develop below is designed for use with event history data (dates of criminal actions for large numbers of individuals).

Most dynamic models of behavior imply restrictions as to the form of the conditional hazard rate function. This is also the case for models of criminal behavior. For example, a popular model of the criminal career assumes that individuals commit crimes at some constant rate λ over the course of a criminal career $(0, T^*]$, where T^* is random. This implies the duration times between successive criminal acts over the period $(0, T^*]$ are distributed exponentially with parameter λ. As we have already discussed, the exponential distribution exhibits no duration dependence. An individual is equally likely to commit a crime in the next small interval of time no matter how long it has been since his last criminal action. Alternative models of criminal activity would not be consistent with an exponential distribution of times between successive crimes. For example, if the opportunity costs associated with committing a crime increased in the length of time since the last crime was committed, while the distribution of potential rewards from criminal actions was constant, the duration distribution of intervals between crimes would exhibit negative duration dependence--the greater the duration since the last crime, the lower the instantaneous rate of committing a crime.

The flexible econometric model presented in Flinn and Heckman (1982a) controls for observed and unobserved heterogeneity in the population by parameterizing the hazard rate function in a general way. If we assume that spell lengths for an individual are i.i.d. conditional on observed and unobserved heterogeneity, and that only one spell is observed for each individual (for notational simplicity), we may write the hazard function as

$$(15) \qquad h_i(w) = \exp\{\underset{\sim}{Z}_i(w)\underset{\sim}{\beta} + \underset{\sim}{A}(w)\underset{\sim}{\gamma} + V_i(w)\},$$

where we have assumed for notational simplicity that the start of the observational period corresponds to calendar time 0. The vector of observable, exogenous individual characteristics at time w is denoted $\underset{\sim}{Z}_i(w)$, and $\underset{\sim}{\beta}$ is a conformable parameter vector. The vector $\underset{\sim}{A}(w)$ consists of polynomial terms in duration, that is, $\underset{\sim}{A}(w) = (w, w^2, \ldots, w^k)$, and $\underset{\sim}{\gamma}$ is a k dimensional parameter vector. An unobserved variable $V_i(w)$ is permitted to be a function of duration. By exponentiating the term in brackets we ensure that $h_i(w)$ is nonnegative as is required, since $h_i(w)$ is a conditional density function.

Many stochastic models of the duration between crimes can be nested within this model as special cases. In many models the role of individual-specific, unobserved heterogeneity is stressed--the $V_i(w)$ in equation 15. Conditional on $V_i(w)$, these models typically restrict $\tilde{\gamma}$ to be a zero vector, thus they posit no duration dependence. Where duration dependence is allowed, functional forms are estimated that restrict the hazard rate function to be monotonically increasing or decreasing in time since the last criminal event. By using our polynomial "approximation," $\exp(\tilde{A}(w)\tilde{\gamma})$, we allow for nonmonotonic patterns of duration dependence. In the absence of a behavioral model that gives the analyst a strong reason to restrict his or her attention to special cases, it can be argued that as general a form of estimating equation as is feasible should be employed. Computationally it is straightforward to introduce the term $\exp(\tilde{A}(w)\tilde{\gamma})$, and we do so in what follows.

The one state renewal model can be generalized in several ways which may prove useful in the study of criminal careers. The assumption of the criminal career being a conditional renewal process (i.e., conditional on other exogenous stochastic processes) can be dropped. Flinn and Heckman (1982b) discuss several forms of departures from the basic renewal process which may be relevant for the analysis of dynamic behavior.

First consider a case in which criminals acquire crime-specific human capital in the course of engaging in criminal behavior. Experienced criminals may be better at avoiding detection, or identifying profitable targets, than nonexperienced criminals. Then if the rewards from legitimate market activity remain approximately constant over the life cycle, we would expect both the frequency with which crimes are committed and the yield from criminal activity to change over the career. We should unambiguously expect the yields from crime to increase; the frequency with which crimes are committed may increase or decrease as criminal experience is acquired. Even if it were possible to measure criminal human capital or yields from crime sufficiently precisely, by conditioning on these characteristics the criminal career could still not be considered a renewal process, since the level of these characteristics depends on the past history of the process.

We can model this departure in a relatively straightforward way. Consider the intervals between crimes for an individual who has committed n crimes. Conditional on all observable exogenous characteristics, we may consider the durations $w_1, w_2, \ldots, w_n$ to be independently but <u>not</u> identically distributed. Then

$$F(w_1, w_2, \ldots, w_n) = \prod_{i=1}^{n} F_i(w_i),$$

but it is <u>not</u> the case that $F_1 = F_2 = \ldots = F_n$. Consider a multiple spell version of (15). Let j index the serial order of the spell (j=1 corresponds to the spell beginning at time 0 and ending with the first crime, j=2 is the spell between the first and second crimes, and so on). Then we can write the hazard rate function for interval j for individual i as

$$(16) \qquad h_{ij}(w) = \exp\{\underset{\sim}{Z}_i(\tau_{ij} + w)\underset{\sim}{\beta}_j + \underset{\sim}{A}(w)\underset{\sim}{\gamma}_j + V_{ij}(\tau_{ij} + w)\},$$

where τ_{ij} is the calendar time at which individual i committed his jth crime, $\underset{\sim}{\beta}_j$, $\underset{\sim}{\gamma}_j$, are parameters associated with the hazard rate function for the jth spell, and V_{ij} is the unobserved heterogeneity component associated with the jth spell for individual i. By analogy with the variance components model often used in the analysis of discrete time panel data, we write

$$V_{ij}(\tau_{ij} + w) = \phi_i + \eta_{ij} + \varepsilon(\tau_{ij} + w),$$

where ϕ_i is an individual specific, spell and time invariant heterogeneity component, η_{ij} is a spell specific, time invariant heterogeneity component, and $\varepsilon(t)$ is white noise (that is, $\varepsilon(t) - \varepsilon(s)$ is normally distributed with mean 0 and variance $(t - s)$ for $t > s$).

In what follows, we neglect continuously varying components of unobserved heterogeneity. While it would be highly desirable to explicitly model such components, their inclusion in the econometric model does not seem computationally feasible. We assume that unobserved heterogeneity components are constant <u>within</u> spells, i.e., $V_{ij}(\tau_{ij} + w) = V_{ij}$. To simplify calculations further, we adopt a one-factor specification of unobserved heterogeneity

$$V_{ij} = C_j\phi_i. \qquad j = 1, \ldots, J,$$

where the C_j are parameters of the model and J is the maximum number of spells observed in the sample. Thus individual heterogeneity is constant over time and spells, though the relationship between ϕ_i and the rate of exit from the spell depends on the serial order of the spell through the parameter C_j.

The rate of criminal activity will not in general depend only on the length of time since the previous crime was committed, but also on the individual's age, and, more important, his previous record of crime commission. Consider spell j. The previous history of individual i's criminal career consists of $\{w_{i1}, w_{i2}, \ldots, w_{i,j-1}; Z(t), 0 \leq t \leq \tau_{i,j-1}\}$. Suppose certain characteristics of this history are of interest to us, for example, the mean, variance, or some other moments of the sample distribution of $\{w_{i1}, w_{i2}, \ldots, w_{i,j-1}\}$. These characteristics are simply functions of the history, $S(H_i(\tau_{i,j-1}))$, where $H_i(\tau_{i,j-1})$ is individual i's history up to time $\tau_{i,j-1}$. Then we can estimate the conditional hazard rate function for the jth interval as

$$h_{ij}(w) = \exp\{\underset{\sim}{Z}_i(\tau_{ij} + w)\underset{\sim}{\beta}_j + \underset{\sim}{A}(w)\underset{\sim}{\gamma}_j + S(H_i(\tau_{i,j-1}))\,\underset{\sim}{\xi}_j + V_{ij}\},$$

where $\underset{\sim}{\xi}_j$ is the parameter vector associated with characteristics of the history up through crime j-1. In this version of the model, spells between crimes are neither identically nor independently distributed, thus the criminal career is modelled as a point process rather than a strict renewal process. Because the process evolves undirectionally in time, the time dependence is recursive. Presumably a model along these lines is required to assess the degree of state dependence in criminal careers--that is, the extent to which the current commission rate depends on the previous criminal history after conditioning on both observed and unobserved exogenous processes.

Up to this point we have assumed that only one type of crime is committed in the population, or at the least, that each individual only commits one type of crime, although different individuals may specialize in different crimes. It is relatively straightforward to generalize the econometric model presented above to cover the possibility of crime switching when each individual may commit any one of a number of types of crimes. Say there are K types of crime, $K > 1$. We will initially restrict our attention to (conditional) renewal processes. Imagine an individual commits a crime of type k at time τ. Then we are interested in estimating the parameters of the length of time between the commission of a type k crime and the commission of all other crimes, for $k = 1, 2, \ldots, K$. For simplicity, assume $k = 2$. At time τ a type 1 crime is committed. The "latent" time to commission of another type 1 crime will be denoted t_{11}^*. The density of these latent times is assumed to exist and be given by $g_{11}(t_{11}^*)$. If type 2 crimes did not exist, this density could be directly estimated using observed durations between successive type 1 crimes. Now denote the "latent" duration time between type 1 crimes and type 2 crimes by t_{12}^* and its associated density by $g_{12}(t_{12}^*)$. It is necessary to assume that the random variables t_{11}^* and t_{12}^* are <u>independent</u>. Now in terms of the <u>observed</u> outcome of the criminal process, a type 1 crime will be the next observed if $t_{11}^* = \min(t_{11}^*, t_{12}^*)$, and a type 2 crime will be observed if $t_{12}^* = \min(t_{11}^*, t_{12}^*)$. Then if $t_{1j}^* = \min(t_{11}^*, t_{12}^*)$, we will observe a type j crime at time $\tau + t_{1j}^*$. Similarly conditional on a type 2 crime at time τ, there will exist latent duration densities $g_{21}(t_{21}^*)$ and $g_{22}(t_{22}^*)$ generating times until the next crime, so $t_{2j}^* = \min(t_{21}^*, t_{22}^*)$. Then in this two-crime world, we would be interested in estimating the parameters of the four latent densities g_{11}, g_{12}, g_{21}, and g_{22}. These densities constitute a complete description of the criminal history.

For the general K state case, we will need a total of K^2 latent density functions to describe the crime process g_{ij}, i,j=1, ..., K. (In addition, we would need to estimate densities g_{0j}, j = 1, ..., K, which correspond to the latent duration densities from initial entry into the population at risk of committing a crime, which we will denote by state 0, until a crime of type j is committed). For each latent density g_{ij}, i = 0, 1, ..., K; j = 1, ..., K, there is a corresponding hazard rate function h_{ij}. The joint density of the k latent durations is given by

$$\prod_{j=1}^{K} h_{ij}(t_{ij}^*) \exp\{-\int_0^{t_{ij}^*} h_{ij}(u)\, du\}, \qquad i = 1, \ldots, K.$$

An individual is observed to commit a type j' crime after the type i crime if the latent time $t_{ij'}^*$ is smallest of the K possible latent times, t_{i1}^*, ..., t_{iK}^*. Let the probability that an individual commits a type j' crime after a type i crime be denoted $P_{ij'}$. Then

$$P_{ij'} = \int_0^\infty [\int_{t_{ij'}^*}^\infty \cdots \int_{t_{ij'}^*}^\infty \{\prod_{j\neq j'}^{K} h_{ij}(t_{ij}) \exp[-\int_0^{t_{ij}^*} h_{ij}(u)du]dt_{ij}^*\}$$

$$\times \{h_{ij'}(t_{ij'}^*)\exp[-\int_0^{t_{ij'}^*} h_{ij'}(u)du]\}]\, dt_{ij'}^*$$

$$= \int_0^\infty h_{ij'}(t_{ij'}^*)\exp\{ - \int_0^{t_{ij'}^*}[\sum_{k=1}^{K} h_{ik}(u)]\, du\}\, dt_{ij'}^*.$$

The conditional density of exit times from state i into state j' given that $t^*_{ij'} < t^*_{ij} (\forall j:\ j \neq j')$ is

$$g(t^*_{ij'} \mid t^*_{ij'} < t^*_{ij}) \quad \forall j:\ j \neq j';$$

$$= \frac{h_{ij'}(t^*_{ij'}) \exp\{ -\int_0^{t^*_{ij'}} [\sum_{k=1}^{K} h_{ik}(u)]\ du \}}{P_{ij'}}.$$

It follows that the marginal density of exit times from state i can be written

$$g_{i\cdot}(t^*_{i\cdot}) = \sum_{j'=1}^{K} P_{ij'} g(t^*_{ij'} \mid t^*_{ij'} < t^*_{ij}) \ (\forall j:\ j \neq j')$$

$$= [\sum_{k=1}^{K} h_{ik}(t^*_{i\cdot})] \exp \{ -\int_0^{t^*_{i\cdot}} [\sum_{k=1}^{K} h_{ik}(u)]\ du \}.$$

The probability that the spell is not complete by some time T, where T is the end of the observation period, is prob $(t^*_{i.} > T) \equiv 1 - G_{i.}(T)$, where $G_{i.}$ is the cumulative distribution function associated with $g_{i.}$. This expression is

$$\text{Prob}(t^*_{i\cdot} > T) = \int_T^{\infty} g_{i\cdot}(t^*_{i\cdot}) dt^*_i$$

$$= \exp \{ -\int_0^T [\sum_{k=1}^{K} h_{ik}(u)]\ du \}.$$

This term enters the likelihood function for incomplete spells at least T in length.

Say we have access to event history data for I individuals. For a given individual i, we observe his criminal career from time of entry into the criminal process, $\tau_0(i)$, until some termination time T(i), which corresponds to the end of the sample period or the time of death (both events are assumed unrelated to criminal activity). In general we observed a total of m(i) crimes over the sample period. Denote the calendar time of each criminal event by $\tau_\ell(i)$, $\ell = 1, 2, \ldots, m(i)$. Now define a function of $s(\tau_\ell(i)) \equiv s_\ell(i)$, which gives the type of crime committed at calendar time $\tau_\ell(i)$.[5] Then conditional on a set of unknown parameters, $\underset{\sim}{\Omega}$, and unobserved person specific heterogeneity component V_i, the likelihood of observing the recorded criminal history for individual i is

$$\pi_i(\Omega \mid V_i) = \Big\{ \prod_{\ell=0}^{m(i)-1} g_{s_\ell(i)s_{\ell+1}(i)}\Big(t^*_{s_\ell(i)s_{\ell+1}(i)} \Big| t^*_{s_\ell(i)s_{\ell+1}(i)} < t^*_{s_\ell(i)j};\ j = 1, \ldots, s_{\ell+1}(i)-1, s_{\ell+1}(i)+1, \ldots, K;\ V_i\Big) \times P_{s_\ell(i)s_{\ell+1}(i)}(V_i)\Big\}\ G_{s_m(i)} \cdot (T(i) - \tau_m(i) \mid V_i),$$

where $t^*_{s_\ell(i)\,s_{\ell+1}(i)} = \tau_{\ell+1}(i) - \tau_\ell(i)$. By substitution,

$$\pi_i(\Omega \mid V_i) = \{ \prod_{\ell=0}^{m(i)-1} h_{s_\ell(i)s_{\ell+1}(i)}\left(t^*_{s_\ell(i)s_{\ell+1}(i)} \middle| V_i\right)$$

$$\times \exp\left[-\int_0^{t^*_{s_\ell(i)s_{\ell+1}(i)}} \left(\sum_{j=1}^{k} h_{s_\ell(i)j}(u \mid V_i)\right) du\right]\}$$

$$\times \exp\left[-\int_0^{T(i) - \tau_{s_m(i)}} \left(\sum_{j=1}^{k} h_{s_m(i)j}(u \mid V_i)\right) du\right]\}.$$

This is the conditional likelihood for an individual observation given a value of the unobserved heterogeneity component V_i--which recall has the substantive interpretation of an individual's inherent propensity to commit crimes. Individual propensities to commit crimes are assumed to be distributed according to $B(V; \underset{\sim}{\Phi})$ in the population, where $\underset{\sim}{\Phi}$ is a vector of parameters which describe the distribution. Note that we assume that V_i is distributed independently of other observable characteristics $\underset{\sim}{Z}$. Then the unconditional likelihood, or integrated likelihood, for an individual observation is

$$\pi_i(\underset{\sim}{\Omega}, \underset{\sim}{\Phi}) = \int \pi_i(\underset{\sim}{\Omega} \mid V_i) dB(V_i; \underset{\sim}{\Phi}).$$

The log likelihood for the entire sample is

$$L(\underset{\sim}{\Omega}, \underset{\sim}{\Phi}) = \ell n \prod_{i=1}^{I} \pi_i(\underset{\sim}{\Omega}, \underset{\sim}{\Phi})$$

$$= \sum_{i=1}^{I} \ell n\, \pi_i(\underset{\sim}{\Omega}, \underset{\sim}{\Phi}).$$

Then the maximum likelihood estimates of $\underset{\sim}{\Omega}$ and $\underset{\sim}{\phi}$ can be obtained under standard regularity conditions as the solution to

$$(17) \quad \max_{\underset{\sim}{\Omega},\underset{\sim}{\phi}} L(\underset{\sim}{\Omega}, \underset{\sim}{\Phi}).$$

Given that the distributional assumptions regarding h and B are correct, the maximum likelihood estimator defined by (17) has optimal statistical properties asymptotically (as the number of individuals and/or the length of individual observation periods grows large).

This model is relatively general and has been used to estimate the stochastic structure of labor market attachments. I hasten to add that the generality of the model seems to preclude treatment of complicated initial-conditions problems or common forms of sample selection. The solutions to these problems seem only to be tractable when sufficient stationarity is imposed--as when the underlying crime process is exponential--see for example Ralph et al. (1981). The difficult choice for the analyst appears to be either to use relatively general econometric models, which require a type and quality of data rarely available to students of criminal behavior, or to tailor our econometric models to the data currently available. This latter option results in stationarity assumptions that are not consistent with the spirit of the dynamic behavioral models in Section 1, and, more important, are not testable. I believe it is essential to first estimate general models for stochastic processes on some "ideal" data set (no doubt yet to be collected) so that we may determine what types of stationary assumptions are reasonable.

Until that time, we should remain cautious in interpreting the results from the empirical analyses of criminal careers.

4. CONCLUSION

In this paper we have presented the approaches to the modelling of criminal careers. In Section 2, a number of dynamic behavioral models of criminal activity were developed, and characteristics of the solutions were discussed. Although closed form solutions are not typically available for dynamic optimization models, numerical methods may be used in a relatively straightforward way. I hope to implement some elaborated versions of these models in future empirical work.

The behavioral models were designed to illustrate the fact that the effect of current choices on future options has potentially important deterrence effects. Thus the fact that an individual facing a one-year sentence if caught committing a crime will face stiffer sentences in the future if caught committing additional crimes will in general affect criminal behavior at all points over the life cycle. The static models usually employed in empirical research are not capable of capturing these _dynamic_ deterrence effects. It was also shown that personal characteristics such as race, age, or drug usage may not be simple indicators of an individual's "inherent" propensity to commit criminal acts, but instead may merely reflect the relative rewards to criminal versus noncriminal actions that the individual faces. Thus these characteristics may be better thought of as indicators of differences in choice sets than differences in preferences. While these interpretations may seem indistinguishable for purposes of conducting empirical analysis, they imply very different policy actions in dealing with criminal behavior.

In Section 3, econometric models of the duration of time between criminal activities (differentiated by type) were presented. These models are capable of capturing the dynamics of the criminal career more adequately than the behavioral models from a strictly empirical perspective. One is left with the difficulty of substantive interpretation of parameter estimates, however, since no explicit behavioral model is used to generate the function estimated. It should be possible to learn something interesting, even if descriptive, about the dynamics of criminal careers from the estimation of such models.

NOTES

[1]Alternatively, it is assumed that differences may be captured in some simple, parametric manner.

[2]In our legal system, individuals charged with crimes are "punished" when found guilty at least partially because the commission of the crime is held to have been an outcome of conscious choice. Only when individuals are adjudicated to have been noncompetent at the time of the crime are they not held legally responsible for the crime they are found guilty of committing. Thus rationality only requires that individuals make consistent choices with respect to some objective and given the choice sets they face. It is a large leap from the assumption of rationality *per se* to the simple utility maximization models developed below.
Unfortunately, it is often the case that discussions of the manner in which criminal behavior should be modelled conclude with the claim that rational choice models are too simplistic to be useful. The point is not whether rationality is a reasonable assumption; I don't believe any social science investigation can be attempted without it. The correct point is that current attempts at behavioral modelling of criminal behavior using the expected utility maximization principle are unquestionably over-simplistic. Realistically, to adequately capture the dynamics of criminal behavior, structural models will have to evolve substantially.

[3]Explicitly incorporating finiteness of life would considerably complicate the analysis while leaving the substantive results unchanged.

[4]This condition is strictly correct only if the wage sequence w_1, w_2, ... is increasing, which is the case in all models we consider.

[5]For example, say that robbery is defined as type 2. If the first and third crimes the individual committed were robberies, then $s_1(i) = s_3(i) = 2$.

REFERENCES

Becker, G.
1975. Human Capital, Second Edition. New York: Columbia University Press.

Bentham, J.
1780. An Introduction to the Principles of Morals and Legislation, J. Burns and H. Hart, eds. London: The Athlone Press (1970).

Blumstein, A., Cohen, J., and Nagin, D., eds.
1978. Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates. Washington, D.C.: National Academy of Science.

Chaiken, J. and Chaiken, M.
1981. Varieties of Criminal Behavior. Santa Monica, Calif.: Rand Corporation.

Flinn, C. and Heckman, J.
1982(a). "Models for the Analysis of Labor Force Dynamics," in Advances in Econometrics I, R. Basmann and G. Rhodes, eds. Greenwich, Conn.: JAI Press.

Flinn, C. and Heckman, J.
1982(b). "New Methods for Analyzing Individual Event Histories," in Sociological Methodology 1982, S. Leinhardt, ed. San Francisco: Jossey-Bass.

Marschak, J.
1953. "Economic Measurements for Policy and Prediction," Chapter 1 in Studies in Econometric Method, Cowles Commission Monograph 14, W. C. Hood and T. C. Koopmans, eds. New York: John Wiley and Sons.

Ralph, J., Chaiken, J., and Houchens, R.
1981. Methods for Estimating Crime Rates of Individuals, Rand Corporation, Santa Monica, Ca.

Schmidt, P. and Witte, A.
1984. An Economic Analysis of Crime and Justice. Orlando, Florida, Academic Press.