Rochester Center for

Economic Research

An Economic Analysis of the Age-Crime Profile

Leung, Siu Fai

Working Paper No. 321
April 1992

University of

Rochester

An Economic Analysis of the Age–Crime Profile

Siu Fai Leung

Rochester Center for Economic Research
Working Paper No. 321

An Economic Analysis of the Age-Crime Profile

Siu Fai Leung

University of Rochester

This paper provides an economic analysis of the relationship between age and crime, an important subject that has completely been ignored in economics. The age-crime profile, which is one of the most established and prominent empirical regularities in criminology, is the focus of the investigation. An economic dynamic model of criminal behavior is formulated to explain the age-crime profile. The model offers a new interpretation of the age-crime profile that is entirely different from the prevailing one in the literature. The implications of the model are testable and are supported by the empirical evidence provided in the paper. The analysis shows that criminal decision can be modeled as a dynamic consumption-investment problem.

First Draft: October 1989
Current Version: April 1992

Department of Economics
Harkness Hall
University of Rochester
Rochester, New York 14627

Shihti Yu provided able research assistance for the numerical analysis in Section III of the paper. I thank Gary Becker, Mark Gritz, Harold Votey, as well as seminar participants at the University of Rochester and the 1991 Western Economics Association International Conference (Seattle, Washington) for their comments.

I. Introduction

In the past few decades, economists have made many significant advances in the analyses of criminal behavior and criminal justice systems. Even skeptical observers have to concede that "It has been a long time since any new idea has been advanced in criminology with as much vigor and conviction as the economic approach to crime" (Messinger and Bittner, 1979, p.21). Despite the admirable success, economists have continued to ignore an important aspect of criminal behavior which has been extensively studied in sociology and criminology —— the relation between age and crime. The age-crime relation is an important subject not only because it has direct implications on the effectiveness of crime control policies such as rehabilitation and selective incapacitation, but also because its complexity poses serious conceptual and empirical challenges to social scientists.

The most well-known finding in the age-crime literature is the age-crime curve (or the age distribution of crime) discovered by Quetelet (1984) more than 150 years ago. Based on some French data, Quetelet found that the propensity for crime (the percentage of people arrested in each age group) rose to a maximum at about age 25 and then declined afterward (see Figure 1). He examined many other causes of crimes (e.g., education, income, seasons) and argued that none could exert such a strong effect on the rise and fall of the propensity of crime. The distinctive age-crime curve led him to conclude that "Among <u>all the causes</u> which have an influence for developing or halting the propensity for crime, the most vigorous is, without contradiction, age." (Quetelet, 1984, p.54, emphasis added)

Quetelet's discovery of the age-crime curve and his emphasis on the effect of age on the propensity for crime have stimulated an enormous amount of work on age and crime in sociology and criminology.[1] The most striking finding in the literature is that the shapes of the age-crime curves of different places and countries at various times are in general parallel to the one Quetelet discovered. The age-crime curve is undoubtedly one of the most firmly established empirical regularities in criminology. Nevertheless, the whole literature is still completely ignored in economics.[2]

Establishing stylized facts and formulating economic models to explain stylized facts have been regarded as two principal goals in economics. In microeconomics, much work has been done using economic theory to explain several stylized facts about lifecycle behavior such as the age-earnings profile, age-wage profile, and age-hours-of-work profile (Killingsworth, 1983). Can economic theory explain the stylized facts about the age-crime profile? If, as Stigler (1970) argues, the details of occupational choice in illegal activity are not different from those encountered in legal activity, then the age-crime profile should be explainable in much the same way as the

[1] Aside from the age-crime relation, Quetelet had made some other contributions to the analysis of human behavior. In fact, some scholars argue that Quetelet should be regarded as the true founder of sociology and criminology, instead of the traditionally acclaimed founders Auguste Comte and Caesar Lombroso. See the preface in Quetelet (1984) for a discussion of his role in the development of criminology and sociology. Quetelet had also made some contributions to statistics, see Stigler (1986) for a detailed account and appraisal.

[2] As aptly pointed out by Winship and Rosen (1988), one of the main contributions of sociological research to economics is to provide empirical facts and phenomena that can inform economic theory. The empirical regularities concerning the age-crime curve that criminologists have carefully established present a serious challenge to economists, who in turn should be able to contribute by providing conceptual insights for the problem. However, it is surprising that the age-crime curve, which can be treated as a lifecycle criminal behavior problem, has not yet been addressed in economics at all. This is in sharp contrast with the extensive ways that other lifecycle problems (such as lifecycle earnings and labor supply) have been analyzed both theoretically and empirically in economics.

age-hours-of-work profile. Nevertheless, some complications arise because of several subtle differences between the age-crime profile and the other age profiles. The age-crime profile is obtained from arrest data (official records or self-reports), therefore each observation denotes a discrete event (arrest). On the other hand, the earnings, wage, or hours of work in the other age profiles measure the magnitude of some continuous variates which are clearly different from the event data. In addition, each arrest is a stochastic event as there are many uncertainties involved in the detection and arrest process. Hence, the arrest age, which measures the time of occurrence of an event, is actually a type of duration data. Because of the intricate differences in the nature of the data, conventional methods used in the analysis of the other age profiles cannot be applied or modified to study the age-crime profile, let alone the economic explanations. A different set of economic tools, namely duration analysis, is required to deal with the age-crime profile.[3] This important distinction has never been recognized in the economic literature and the aim of this paper is to provide a first systematic attempt to tackle the problem.

Apart from bringing forward a new and important subject to economists' attention, the paper makes two main contributions. First, it demonstrates that economic theory via duration analysis can rigorously and fruitfully be applied to analyze the age-crime profile. The theory is testable and is supported by the empirical evidence examined. Second, the theory offers a new interpretation of the age-crime profile which is entirely different from

---

[3] See Heckman and Singer (1986), Kiefer (1988), or Pudney (1989) for a survey of economic duration analysis. While the literature on duration analysis is mostly concerned with econometric issues, the focus of the present analysis is on economic modeling.

the prevailing one in the literature. The policy implications are accordingly different from those derived from the prevailing interpretation.

The rest of the paper is organized as follows. Section II reviews briefly the age-crime literature, describes the focus of the paper, and introduces the prevailing interpretation of the age-crime profile. By using some basic concepts in duration analysis, Section III develops an economic dynamic model of crime to explain the age-crime profile. A numerical example is used to demonstrate that the model can generate the age-crime profile. A comparison between the implications of the model and the prevailing interpretation is made. Section IV provides three sets of empirical evidence to support the implications of the model. Section V concludes the paper.

II. The Age-Crime Profile and the Prevailing Interpretation

Ever since Quetelet's pioneering work, many researchers have analyzed data from different countries at various times and have generally found the same age-crime profile: the crime rate (arrest rate) increases rapidly from the juvenile years to reach a maximum in the late teens and then steadily declines afterward.[4] The two common features of these age-crime profiles are that they are unimodal and peak around the teenage years. These studies also find that the age-crime profiles do not only have a remarkable resemblance for different countries in different times, but they are also very similar under a variety of classifications: first convictions, different crimes (such as burglary, fraud), different categories of crime (violent versus

[4] See, e.g., England and Wales in 1842-44 (Neison, 1857); England in 1908 (Goring, 1913); England in 1938, 1961, 1983 (Farrington, 1986); Argentina in the 1960s (DeFleur, 1970); United States in 1982 (Farrington, 1986); United States in 1940, 1960, 1980 (Steffensmeier et al. 1989). More references can be found in Hirschi and Gottfredson (1983).

non-violent), different sex (males versus females), and different races (white versus nonwhite).[5] In general, the unimodal feature appears persistently in most age-crime profiles, although the ages at which the curves peak may differ. Based on all these findings, one can conclude that unimodality and positive skewness are the two salient features that characterize the age-crime profile. To illustrate these findings, Figures 2 and 3 plot the age-crime profiles for two types of classifications: crime types and sex. These curves are derived from the data reported in Quetelet (1984) and details on the derivations are reported in Appendix A. Figure 2 shows that the age-crime profiles for crimes against persons and crimes against property are similar; whereas Figure 3 shows that the age-crime profile of men resembles that of women. Parallel to the one in Figure 1, the age-crime profiles in Figures 2 and 3 are in general unimodal and positively skewed. The peaks occur slightly after the teenage years.

The findings on the similarity of the age-crime profiles in the literature are striking especially when one takes into account the fact that the population arrest rates have been increasing over time in many countries. For instance, while the population arrest rate of the U.S. in 1976 almost doubles that of 1965, Blumstein and Cohen (1979) find that the age-crime profiles in 1965 and 1976 are almost identical.

Since the age-crime profiles are usually obtained from aggregate cross-sectional data (synthetic cohorts), the age-crime relation is likely to be the outcome of the interaction of three different effects: cohort, period, and age. In order to study the relation between age and crime from aggregate

---

[5] See, e.g., Hirschi and Gottfredson (1983), Farrington (1986), and Steffensmeier et al. (1989).

data, the cohort effect and the period effect have to be separated from the age effect. Using FBI data from 1964 to 1979 and assuming that there is no period effect, Greenberg (1983) analyzes the age-cohort data and finds that the age effect is still present, but it is less pronounced than the one obtained from cross-sectional data. In contrast, using England and Wales data from 1961 to 1983, Farrington (1986) finds that the age effect is parallel to the one obtained from cross-sectional data even after controlling for cohort and period effects. To avoid the problems of separating the three effects in cross-sectional data, some researchers turn to longitudinal data to study the age-crime relation. As summarized in Farrington (1986), the general finding is that the age-crime profiles mirror the familiar pattern obtained from cross-sectional data. Given these results, it is not surprising that the age-crime profile has become one of the most widely accepted empirical regularities in sociology and criminology. However, the interpretation of the profile is far from settled.[6]

The age-crime profile describes the age distribution of the aggregate arrest rate (the proportion of people arrested in each age group). In each age group, the aggregate arrest rate is determined by the proportion of

---

[6] An example is the recent controversy sparked by Hirschi and Gottfredson (1983). Since the age-crime profiles for different places, times, sex, and crime types are similar, Hirschi and Gottfredson (1983) contend that the relation between age and crime is invariant. Therefore, the age-crime relation holds independently of, and cannot be explained by reference to, other physical or social factors. That is, age has a direct causal effect on crime. This provocative theory has generated an ongoing intense debate in the literature. See, e.g., Greenberg (1985), Hirschi and Gottfredson (1985), Farrington (1986), Gottfredson and Hirschi (1986), Blumstein, Cohen and Farrington (1988a), Gottfredson and Hirschi (1988), Blumstein, Cohen and Farrington (1988b), Tittle (1988), and Steffensmeier et al. (1989).

people engaged in criminal activities and their intensities of offending. Hence, the age-crime profile may be generated by changes in the participation rate or changes in the intensity rate with age (or both).[7] Of course, law enforcement agencies may also play a role in generating the age-crime profile. For example, assume every age group has the same participation rate and the same intensity rate of offending. If police officials deliberately spend more resources in detecting and apprehending young offenders than old offenders, this will also create the age-crime profile. Nevertheless, this does not appear to be a major determinant of the age-crime profile because it is difficult to explain why the age-crime profile is so pervasive and consistent over space and time, unless different countries in different times employ the same policy of targeting young offenders. It is highly unlikely that the age-crime profile is an artifact of the criminal justice system.

Whether the age-crime profile is generated by changes in the participation rate or changes in the intensity rate of offending (or both) is largely an empirical question that cannot be settled a priori. The empirical evidence to date shows that both participation and intensity play a role in generating the profile (Loeber and Synder, 1990). Thus, the challenge is to explain how changes in the participation rate and the intensity rate with age can create the age-crime profile. Most previous work considers only the participation aspect. Therefore, this paper will focus on the relatively neglected and conceptually more difficult problem of analyzing the intensity aspect. Specifically, the objective is to study

[7] The distinction between participation and intensity is familiar in economics. For example, aggregate labor supply is determined by the number of workers (participation) and their hours of work (intensity).

whether changes in the intensity rate alone can generate the age-crime profile. To achieve this, it is necessary to assume in the following analysis that the participation rate does not vary with age, so that the age-crime profile is solely derived from changes in the intensity rate with age. This working assumption is made in order to isolate the intensity effect from the participation effect. Thus, given that an offender has made the participation decision, the problems are to investigate:

(P1) how the intensity rate of offending changes with time (age), and

(P2) whether the time path of the intensity rate can generate the age-crime profile.

In other words, the problems are to determine the age-intensity profile and study its relation with the age-crime profile.

The prevailing interpretation of the age-crime profile in the criminology literature offers an answer to these two problems. A clear statement can be found in the influential report compiled by the National Research Council's Panel on Research on Criminal Careers (Blumstein et al. 1986).[8] The Panel contends that:

> The distinctive age patterns in aggregate measures may be due either to changes in participation or in individual frequency rates for active offenders. In the former case, the peak rates of criminal activity would result from growing participation in crime during the late teen years, followed by declining participation as increasing numbers of offenders end their criminal careers. In the latter case, peak rates would arise from variations in the intensity of offending by a fairly fixed group of active offenders, with individuals' frequency rates increasing during the juvenile years and then gradually declining with age (Blumstein et al. 1986, vol. 1, pp.23-24, emphasis added).

---

[8] Gottfredson and Hirschi (1986, 1988) give a critical account of how the prevailing views influence the allocation of federal research funds and dominate the research community in criminology.

This prevailing interpretation claims that if the age-crime profile is generated by variations in the intensity rate, the age-intensity profile must have <u>the same shape</u> as the age-crime profile. This implies that the intensity rate of offending increases rapidly from the juvenile years to reach a maximum in the late teens and then steadily declines afterward. In other words, the shape of the age-crime profile is exactly derived from the shape of the age-intensity profile. This interpretation, which appears natural and plausible, will be evaluated in the next section.

III. An Economic Analysis

In this section, I will develop an economic dynamic model of crime, use it to explain the age-crime profile, and contrast the model's implications with the prevailing interpretation. Before formulating the model, it is necessary to further formalize problems (P1) and (P2) mentioned in the previous section.

An active offender faces an uncertain arrest time since there are many stochastic elements involved in the detection and arrest process. An arrest is a stochastic event and the arrest time t is therefore a nonnegative random variable. Similar to the unemployment spell or strike duration in economic duration analysis (or the failure time in industrial reliability theory and biometric survival analysis), the arrest time is a duration variable. Hence, it is natural to model the underlying behavior using economic duration analysis and hazard functions. Let f(t) and F(t) be respectively the probability density function (pdf) and cumulative distribution function of the arrest time of an active offender. The age-crime profile is related to the pdf f(t) in a special way.

A typical age-crime profile is a histogram describing the frequency distribution of the arrest rate according to age. With a constant participation rate, this implies that the age-crime profile is a statistical estimate of the underlying probability density function of the arrest time of active offenders. Assuming homogeneous offenders, the age-crime profile is therefore an estimate of the pdf $f(t)$. Since the sample sizes in most age-crime studies are usually fairly large, the age-crime profile will approximate $f(t)$ reasonably well.[9] Hence, for practical and analytical purposes, the age-crime profile can be regarded as $f(t)$. Therefore, to explain why the age-crime profile is unimodal and positively skewed is equivalent to explaining why the pdf $f(t)$ has such features. Let $c(t)$ be an offender's intensity rate of offending at time t, problems (P1) and (P2) can now be formalized as: (P1') to study the age-intensity profile $c(t)$, and (P2') to investigate whether it can generate the age-crime profile $f(t)$.

Let $h(c(t))$ be the hazard rate of arrest at time t. The higher the intensity rate of offending, the higher will be the hazard rate of arrest.[10]

[9] The precision of the approximation increases with the sample size. Using standard asymptotic theory, it can be proved that under certain regularity conditions, the histogram (age-crime profile) will converge uniformly to the true probability density function ($f(t)$) as the sample size tends to infinity (see, e.g., Theorem 2.3.1 in Prakasa Rao, 1983).

[10] So far the intensity rate of offending has not yet been exactly defined. It can be defined by the effort (or time) allocated to criminal activity or the frequency of offending. It can also be the intensity (severity) of violation of laws and regulations, e.g., the number of items stolen, the value of properties damaged, or the degree of forgery. Because there are so many different types of crime and different dimensions of involvement in crime, an exact definition of the intensity rate is impossible and is unnecessary for the purpose of a general analysis. It is therefore sufficient for analytical purpose to define the intensity rate in the following general way. The intensity rate, which measures an offender's involvement in criminal activity, is a variable $c(t)$ such that both the hazard rate of arrest and the returns from crime are increasing in $c(t)$.

It is well-known from duration analysis that the pdf $f(t)$ can be expressed in terms of the hazard function and the cumulative distribution function in the following way:

$$f(t) = h(c(t))[1-F(t)]. \tag{1}$$

Equation (1) describes the relation between the age-crime profile $f(t)$ and the age-intensity profile $c(t)$. As will be seen below, this equation forms the basis of the whole analysis. In particular, it is the key to address the problem (P2′).

The dynamic decision problem that a representative offender faces will now be formulated as an optimal control problem. Consider a risk neutral offender with a finite lifetime T.[11] Let $J(\tilde{c},F(x),x)$ be the present value of expected lifetime returns at time x, given the state variable $F(x)$ and the control variable $\tilde{c}$, where $\tilde{c} = \{c(t)\colon t \in [x,T]\}$. Subject to the law of motion (1), the offender's decision problem at time x is to choose $c(t)$, $t \in [x,T]$, to maximize

$$J(\tilde{c},F(x),x) = \int_x^T e^{-r(t-x)}\pi_1(c(t))\left(\frac{1-F(t)}{1-F(x)}\right)dt + \int_x^T e^{-r(t-x)}\cdot$$

$$\left[-\theta(c(t)) + \int_t^{t+s(T-t)} e^{-r(z-t)}\pi_2 dz + \int_{t+s(T-t)}^T e^{-r(z-t)}\pi_3(c(t))dz\right]\left(\frac{f(t)}{1-F(x)}\right)dt. \tag{2}$$

The objective function (2) can be explained as follows. At any time $t > x$, the offender will continue to enjoy the returns from crime $\pi_1(c(t))$, provided that he has not yet been caught, the probability of which is

---

[11] It is easy to generalize the model to handle an uncertain lifetime (i.e., T is a random variable) and all the results will carry through as long as T is a bounded random variable.

$[1-F(t)]/[1-F(x)]$. If he is caught at time t (the instantaneous probability is $f(t)/[1-F(x)]$), then he faces two possible forms of punishment: a fine of $\theta(c(t))$ and a length of incapacitation $s(T-t)$. His returns during incapacitation will be π_2. The term $s(T-t)$ indicates that the length of incapacitation is a fraction s, $s \in [0,1]$, of the remaining lifetime $(T-t)$ of the offender. After the offender is released at time $t+s(T-t)$, his returns will be $\pi_3(c(t))$ from some legal activity. The returns π_3 are allowed to depend on $c(t)$ in order to capture the possibility of a stigma effect, e.g., if having a criminal record lowers an individual's earnings, then the returns π_3 will decrease with the intensity rate of offending $c(t)$.

The model is essentially a three-period model. The offender participates in criminal activity in the first period. If he is caught, he will enter into the second period of incapacitation. After he is released, he will no longer participate in any criminal activity in the last period. For simplicity, recidivism is ignored. It is straightforward to extend the model to deal with recidivism, but this will greatly complicate the analysis and thus will not be pursued here.[12]

Since the age-crime profile is usually derived from arrest data which contain recidivists, is it appropriate to use the three-period model to explain the age-crime profile? There are three ways to answer the question. First, if offenders are myopic, then the three-period model will be relevant. Second, the main difference between repeat offenders and non-repeat offenders should only be on the <u>level</u> of the intensity rate of offending, and <u>not</u> on the overall time path of the intensity rate. This is

[12] A preliminary attempt to deal with recidivistic behavior is available from the author on request.

because, given that both types of offenders engage in crime and have not yet been arrested for the current offenses, the number of previous arrests (positive for repeat offenders and zero for non-repeat offenders) is fixed and does not vary with time. Since the number of previous arrests is a time invariant variable during the time interval between arrests, the time paths of the intensity rate of offending of repeat and non-repeat offenders should be similar, although the levels may be different. Consequently, the overall shapes of the age-crime profiles of repeat and non-repeat offenders should be similar. There is substantial evidence supporting this argument. Consider the age-crime profile for first arrests (or convictions). This type of age-crime profile excludes repeat offenders by restricting the sample only to people arrested (or convicted) for the first time. Many empirical studies in different countries at various times find that the age-crime profiles for first arrests are similar to the general age-crime profile.[13] This evidence suggests that the impact of recidivistic behavior on the age-crime profile is insignificant. Hence, although the three-period model ignores recidivism, it should still be sufficient for the purpose of explaining the essential features of the age-crime profile. Third, the three-period model can at least be used to explain the age-crime profile for first arrests.

The following technical assumptions will be maintained throughout the paper: $\pi_1'' < 0$, $\pi_2' = 0$, $\pi_3' \leq 0$, $\pi_3'' \leq 0$, $\theta' \geq 0$, $\theta'' \geq 0$, $h' \geq 0$, $h'' \geq 0$, $s' \geq 0$, $s'' \geq 0$, $\pi_1(0) = \pi_3(0) \geq 0$, $\theta(0) = 0$, $h(0) = 0$, and $s(0) = 0$. The assumption that π_2 is independent of c implies that the offender's returns

[13] See, e.g., Goring (1913), Christensen (1967), Wolfgang et al. (1972), Wadsworth (1975), Blumstein and Graddy (1982), Farrington (1983), and Fry (1985).

during incapacitation do not depend on the intensity rate of offending before he is caught. This is a reasonable assumption in many situations and help simplify the model. Relaxing this assumption will only slightly complicate the technical details and will not affect the results of the paper. The assumption $\pi_1(0) = \pi_3(0)$ is a natural one, for if an individual chooses $c(t) = 0$, then the returns $\pi_1(0)$ should be equal to the returns $\pi_3(0)$. The last three assumptions state that the fine, the hazard rate of arrest, and the fraction of lifetime incapacitated are zero if the intensity rate of offending is zero. It is straightforward to justify the remaining assumptions on the curvatures of the functions.

The optimization problem (2) is an optimal control problem with $c(.)$ as the control variable and $F(t)$ as the state variable. The objective function (2) can be considerably simplified by some straightforward calculations. For simplicity and without loss of generality, let $x = 0$ and $F(0) = 0$. Appendix B shows that

$$J(\tilde{c},F(0),0) = [1-e^{-rT}]\pi_2/r + \int_0^T e^{-rt}\cdot$$

$$\left\{[\pi_1(c(t))-\pi_2]-\theta(c(t))h(c(t))+D(c(t),t)[\pi_3(c(t))-\pi_2]h(c(t))\right\}[1-F(t)]dt. \quad (3)$$

where $D(c,t) = [e^{-rs(c)(T-t)}-e^{-r(T-t)}]/r$. This simplification of the objective function shows that the offender has at least $[1-e^{-rT}]\pi_2/r$ as the lifetime returns; and $\pi_1-\pi_2$ and $\pi_3-\pi_2$ are the <u>net</u> returns from crime and from legal activity respectively. Without loss of generality, π_2 is normalized to zero; so that (3) can further be simplified to

$$J(\tilde{c},F(0),0) =$$

$$\int_0^T e^{-rt}\left\{\pi_1(c(t))-\theta(c(t))h(c(t))+D(c(t),t)\pi_3(c(t))h(c(t))\right\}[1-F(t)]dt. \quad (4)$$

The term $\pi_1(c(t))-\theta(c(t))h(c(t))$ can be interpreted as the expected net returns from criminal activity at each time instant since $h(c(t))$ is the instantaneous probability of arrest and $\theta(c(t))$ is the fine so that $\theta(c(t))h(c(t))$ is the expected loss. The term $D(c(t),t)$ can be interpreted as a discount factor on the returns $\pi_3(c(t))$. The discount arises from the punishment from incapacitation which postpones the flow of returns $\pi_3(c(t))$ to the time the offender is released.

Let $U(c,t) = \pi_1(c)-\theta(c)h(c)+D(c,t)\pi_3(c)h(c)$, and $\lambda(t)$ be the multiplier for the law of motion (1), then the Hamiltonian for the optimal control problem is given by

$$H(c,F(t),t) = e^{-rt}U(c,t)[1-F(t)] + \lambda(t)h(c)[1-F(t)].$$

Since the objective of the paper is to analyze the intensity decision and not the participation decision, the optimal control will be assumed to be an interior solution. The necessary conditions for optimality for an interior solution are:

$$\partial H(c,F,t)/\partial c = [e^{-rt}U_1(c,t) + \lambda(t)h'(c)](1-F) = 0, \tag{5}$$

$$\partial^2 H(c,F,t)/\partial c^2 = [e^{-rt}U_{11}(c,t) + \lambda(t)h''(c)](1-F) \leq 0, \tag{6}$$

$$\lambda'(t) = -\partial H(c,F,t)/\partial F = e^{-rt}U(c,t) + \lambda(t)h(c), \tag{7}$$

$$\lambda(T) = 0, \tag{8}$$

where (8) is the transversality condition. Since there is a positive probability that the offender will never be caught, i.e., $F(T) < 1$, therefore the necessary condition (5) becomes

$$e^{-rt}U_1(c,t) + \lambda(t)h'(c) = 0. \tag{9}$$

The following proposition characterizes some properties of the multiplier $\lambda(t)$:

Proposition 1: (i) For $t \in [0,T]$, $\lambda(t) = -e^{-rt}[\mathrm{Max}_{\tilde{c}}\ J(\tilde{c},F(t),t)]$. (10)

(ii) For $t \in [0,T)$, $\lambda(t) + e^{-rt}D(c(t),t)\pi_3(c(t)) < 0$. (11)

Proof: See Appendix C.

Proposition 1 shows that $-e^{rt}\lambda(t)$ is the maximized present value of expected lifetime returns at time t. The inequality (11) will be found to be useful in later analysis. To investigate the time path of c(t), I assume here for simplicity that both s and π_3 do not depend on c. This assumption is made in order to avoid technical details which may obscure the essential elements of the model. The results obtained below will carry through in the general case when both s and π_3 are allowed to depend on c; a full treatment is provided in Appendix D. When s is not a function of c(t), D(c(t),t) can just be written as D(t). Let

$$M = e^{-rt}U_{11} + \lambda h'' = e^{-rt}(\pi_1'' - \theta''h - 2\theta'h' - \theta h'') + (\lambda + e^{-rt}D\pi_3)h''.$$

It follows from the assumptions ($\pi_1'' < 0$, $\theta'' \geq 0$, $\theta' \geq 0$, $h' \geq 0$, $h'' \geq 0$) and the inequality (11) that $M < 0$, hence the second-order sufficient condition is satisfied.

At $t = T$, (10) implies $U_1(c(T),T) = 0$, which implies $\pi_1'(c(T)) - \theta'(c(T))h(c(T)) - \theta(c(T))h'(c(T)) = 0$, since $D(T) = 0$. Hence, at the terminal time, the offender will choose the effort that maximizes $\pi_1 - \theta h$, which is intuitively reasonable since the offender cannot be incapacitated after that date. At any $t < T$, $c(t) \leq c(T)$. To see this, (9) can be written as $\pi_1'(c(t)) - \theta'(c(t))h(c(t)) - \theta(c(t))h'(c(t)) = -[D(t)\pi_3 + e^{rt}\lambda(t)]h'(c(t))$. For $t = T$, the right-side is zero; for $t < T$, the right-side is positive (by (11)). This immediately implies that $c(t) \leq c(T)$ for $t \in [0,T]$, because $\pi_1(c) - \theta(c)h(c)$ is strictly concave in c. In other words, the intensity rate at any time is lower than or equal to the intensity rate at the terminal time T.

Differentiate (9) with respect to t, and replace λ' and λ respectively by (7) and (9),

$$\begin{aligned} c'(t) &= e^{-rt}[(r+h)U_1 - U_{12} - Uh']/M \\ &= e^{-rt}Z(c,t)/M, \end{aligned} \tag{12}$$

where

$$\begin{aligned} Z(c,t) = {} & [r+h(c)][\pi_1'(c)-\theta'(c)h(c)-\theta(c)h'(c)] - [\pi_1(c)-\theta(c)h(c)]h'(c) \\ & + (1-s)e^{-rs(T-t)}\pi_3 h'(c). \end{aligned} \tag{13}$$

The sign of $c'(t)$ is the opposite of $Z(c,t)$. The following proposition characterizes the age-intensity profile $c(t)$:

Proposition 2: If s and π_3 do not depend on c, then $c'(t) \geq 0$ for $t \in [0,T]$.

Proof: Suppose $c'(t) < 0$ for some $t \in [0,T]$. Since $c(t) \leq c(T)$ for $t \in [0,T]$, then there must exist t_1 and t_2 ($t_1 < t_2$) such that $c(t_1) = c(t_2)$, and $c'(t_1) < 0$ and $c'(t_2) > 0$. This implies $Z(c(t_1),t_1) > 0$ and $Z(c(t_2),t_2) < 0$. Accordingly, $Z(c(t_1),t_1) - Z(c(t_2),t_2) > 0$. However, $c(t_1) = c(t_2)$ and (13) implies that

$$Z(c(t_1),t_1) - Z(c(t_2),t_2) = (1-s)e^{-rsT}(e^{rst_1} - e^{rst_2})\pi_3 h'(c(t_1)).$$

Since $t_1 < t_2$, therefore $Z(c(t_1),t_1) - Z(c(t_2),t_2) \leq 0$. Contradiction.

Proposition 2 shows that the intensity rate of offending increases with time, which means that the offender will progress toward increasingly more intense and serious criminal involvement. This result still holds when s and π_3 are allowed to depend on c, see Appendix D for a proof. If learning is introduced into the model, the result will likely to be even stronger because the offender will become more skillful and more specialized in criminal activity over time. Since $h'(c) \geq 0$, Proposition 2 also implies that the hazard rate of arrest increases with time, i.e., there is positive duration dependence.

Several factors contribute to the increase in the intensity rate over time. Older offenders have fewer years of life remaining so that they are less concerned about the future consequences of current action. Furthermore, the deterrent effect of punishment diminishes as the offender approaches the terminal time because the length of incapacitation $s(T-t)$ decreases with time t, for any given s and T. In addition, since the probability of not being arrested at any time t is decreasing in t (i.e., $d[1-F(t)]/dt \leq 0$), the expected returns from criminal activity (see (4)) fall over time, other things being equal. A higher intensity rate of offending (and hence higher returns) is necessary to offset the loss in the expected returns. It can be seen that all these factors stem from the fact that the offender has a finite lifetime. The increase in the intensity rate over time is driven by the end-of-horizon effects.[14]

An interesting way to interpret the above results, which also reveals the economics of the model, is to treat the criminal decision problem as a lifetime consumption-investment decision problem. At time zero, every individual is endowed with one unit of capital which cannot be traded or accumulated. The capital is the probability that an individual will not be arrested. In other words, the capital stock at any time t is given by

---

[14] It should be emphasized that the model does not exclude the possibility that an offender may quit the criminal career when he is old (a corner solution). The proposition only states that given participation (an interior solution), the intensity of offending will increase with time.

1-F(t), the probability that an individual will not be arrested by time t. Every individual begins with one unit of capital because at time zero, $1-F(0) = 1$ (since $F(0) = 0$). Although the capital cannot be traded or accumulated, it can be used to produce consumption. If an individual engages in criminal activity, then he begins to reduce the capital stock over time because the probability that he will not be arrested will fall (i.e., 1-F(t) decreases with t). The more intense his involvement in criminal activity, the faster will the capital stock be depleted. Exhausting the capital stock means that the individual can no longer engage in criminal activity. On the other hand, the individual's consumption depends on his intensity of offending. The higher the intensity, the higher will be the consumption. Therefore, a dynamic tradeoff exists between consumption and capital decumulation. As the model reveals, the optimal strategy is to increase the intensity (and hence the consumption) over time. This means that the individual will forego some current consumption in order to sustain a higher consumption in the future. Such a strategy is reasonable because if he chose to have a high consumption in the beginning, then he would exhaust the stock of capital too early so that he would not be able to enjoy any later gains at all, resulting in a lower lifetime expected utility. The model shows that the result still holds even if the individual has a high discount rate on future consumption.[15]

Does the age-intensity profile generate the age-crime profile? This can be checked by totally differentiating (1) with respect to t and replacing f(t) by (1),

---

[15] This differs from the classical result of Yaari's (1964) dynamic consumption-investment model which states that consumption will rise with time if and only if the discount rate is smaller than the interest rate.

$$f'(t) = \{h'(c)c'(t) - [h(c(t))]^2\}[1-F(t)]. \tag{14}$$

The sign of $f'(t)$ is the same as that of $h'c'-h^2$. It is clear that $c' \geq 0$ does not necessarily imply that $f(t)$ is unimodal because the expression $h'c'-h^2$ depends on the functional forms of the hazard function and c'.[16] The other fundamental feature of the age-crime profile, positive skewness, is even more difficult to check because skewness is highly dependent on the functional forms. Without explicitly specifying the functional forms in the model, it is not possible to examine at this level of generality whether the model (or <u>any</u> other model) can generate the age-crime profile.[17] Hence, the analysis has to be proceeded by way of an example.

Example: Let $\pi_1(c) = \pi_3 + w_1c - w_2c^2$, $\theta(c) = w_3c$, $h(c) = w_4c$, $s(c) = s$, where π_3, w_i $(i=1,2,3,4)$, and s are positive real numbers, with $s \in (0,1)$.

Let $w_5 = w_2+w_3w_4$, it is straightforward to verify that

$$c'(t) = (w_4/2)c^2 + rc - (rw_1-\pi_3w_4)/2w_5 - (1-s)e^{-rs(T-t)}\pi_3w_4/2w_5, \tag{15}$$

$$h'c'-h^2 = w_4[-(w_4/2)c^2+rc-(rw_1-\pi_3w_4)/2w_5-(1-s)e^{-rs(T-t)}\pi_3w_4/2w_5]. \tag{16}$$

To examine the sign of $h'c'-h^2$, it is necessary to solve $c(t)$ from (15). However, (15) is a differential equation which does not admit a closed-form

---

[16] If $f(t)$ is unimodal at $t^* \in (0,T)$, then $h'c'-h^2$ will be positive for $t < t^*$ and negative for $t > t^*$.

[17] Notice that in some cases, it is easy to check that a model can <u>never</u> generate the age-crime profile. For example, in Davis' (1988) intertemporal model of crime, the intensity rate of offending does not change with time. This implies that his model can <u>never</u> generate the age-crime profile because $c'(t) = 0$ implies $f'(t) = -h(c)f(t) \leq 0$. Therefore, the pdf $f(t)$ in Davis' model is always decreasing in t. Consequently, $f(t)$ can never have the unimodal shape of the age-crime profile. In fact, $f(t)$ is an exponential density function because $c'(t) = 0$ implies $dh(c(t))/dt = 0$, and the only distribution function with a constant hazard rate is the exponential distribution.

solution because of the presence of the exponential term. Consequently, it is not feasible to examine the sign of $h'c'-h^2$ analytically. Instead, numerical methods have to be employed to solve (15) and hence the age-crime profile can only be generated numerically. Figures 4 and 5 report some simulation results. By choosing $w_1 = 10$, $w_4 = 150$, $w_5 = 4$, $\pi_3 = 0.355$, $s = 0.1$, $r = 0.8$, and $T = 1$, a numerical solution for the differential equation (15) is obtained by means of the Runge-Kutta method (Press et al., 1986).[18] Figure 4 plots the solution c(t) versus t, which indicates clearly that the intensity rate increases with time. By substituting the solution of c(t) into (1), a plot for f(t) is obtained and is displayed in Figure 5. It is clear that f(t) is unimodal, positively skewed, and closely resembles the age-crime profile. Therefore, the example demonstrates that the model can generate the age-crime profile numerically. Although it is impossible to show analytically that the model always generates the age-crime profile, the result that it can do so in some numerical examples is sufficient to establish the conclusion that the age-crime profile can solely be derived from changes in the intensity rate alone.[19]

---

[18] The lifetime $T = 1$ is chosen to indicate that the time (age) scale can be arbitrarily calibrated.

[19] In virtually every theoretical work on age related profiles, the only way to explain the profiles is to use examples and specific functions because theories without functional specifications are usually too general to yield definite results. Clearly, this approach may not be entirely satisfactory. In his analysis of the age-earnings profile, Rosen (1976, p.47) points out that "the theory is far too general to be taken as a serious candidate for accounting for observed behavior. Surely there exists a functional specification ... sufficiently complex to fit the data well. Thus, the practical issue ... is to find a parsimonious form that works tolerably well." Thus, although it may always be possible to find functional specifications sufficiently complex to generate the age-crime profile, the specifications should be parsimonious and work tolerably well. The example used in the text is a very simple one and seems to meet the criteria Rosen suggested.

How does the age-intensity profile generate the age-crime profile? The key lies in the selection effect. Since the hazard rate of arrest increases with the intensity rate of offending, the proportion of old offenders is smaller because a large portion of them had already been apprehended when they were young. Fewer offenders can successfully avoid arrest as they get old because their intensity rates of offending increase over time. The observed age-crime profile indicates more intense selection among the young offenders since it is more difficult to sustain a criminal career involving more serious offenses for an extended period of time.

The above interpretation of the age-crime profile differs substantially from the prevailing interpretation outlined in Section II. The model clearly does not support the prevailing interpretation. It is not necessary for the age-intensity profile to have the same shape as the age-crime profile in order that the former can generate the latter. This can be seen from (14) that $f'(t) > 0$ (< 0) does not imply $c'(t) > 0$ (< 0), because of the additional term $[h(c)]^2$. Consequently, the shape of the age-intensity profile $c(t)$ will not be the same as that of the age-crime profile $f(t)$. On the other hand, the model shows that a montonically increasing intensity rate can generate the age-crime profile.

The difference between the two interpretations is not a minor one because the policy implications differ considerably. The model implies that crime control strategies should focus on identifying and convicting active offenders because if they are not arrested, they will commit progressively more serious crimes over time. On the other hand, the prevailing view suggests that active offenders do not present such a serious problem since their intensity rates of offending will eventually fall with time. Given limited budgets and resources, the prevailing interpretation suggests that law enforcement agencies should pay more attention to teenagers (relative to other age groups) because the intensity of offending peaks around the teens. In contrast, the model recommends that more emphasis be put on the older age groups since their intensities of offending are higher than the teenagers'.

IV. Empirical Evidence

The model developed in the previous sections yields two testable implications: the intensity rate of offending as well as the hazard rate of arrest increases with age. The second implication means that there is positive duration dependence. In contrast, the prevailing interpretation implies that both the intensity rate of offending and the hazard rate of arrest will rise and then fall with age. This section provides three sets of empirical evidence to support the implications of the model.

The first set of evidence comes from findings on the relation between the seriousness of offenses and age. The seminal longitudinal study of a Philadelphia cohort of males born in 1945 by Wolfgang, Figlio and Sellin (1972) find that the average seriousness of offenses, measured by a scale that they design, increases with age. A subsequent follow-up of 10 percent of the original cohort finds that the average seriousness of offenses continues to increase with age (Wolfgang, 1980). Offenses committed during the adult period are significantly more serious than those committed during the juvenile period. A more recent study on arrest data in the Detroit SMSA area by Blumstein et al. (1988) also finds that white offenders on average engage in increasingly serious offenses. There are several other studies which also find a similar increase in offense seriousness over the criminal career, see the survey in Cohen (1986). Since the severity of offenses is an important component of the intensity of offending, these empirical findings support the implication of the model that the intensity rate of offending increases over time.

The second set of evidence is drawn from the literature on the fitted age-crime profiles. By fitting the age-crime profiles using a variety of different probability density functions, Farrington (1986) finds that the gamma density $b^a e^{-bt} t^{a-1}/\Gamma(a)$ resembles the age-crime profile quite well, with $a = 1.8$ and $b = 0.11$ ($\Gamma(.)$ is the gamma function). He obtains a better fit by moving up to a three-parameter probability density function of the form $at^b e^{-ct}$, which includes the gamma density as a special case. For the different sets of parameter values (a,b,c) that fit the data well, b is always found to be strictly positive. What do these distribution functions imply about the shapes of the corresponding hazard functions? Do they imply positive duration dependence?

The best way to answer these questions is to use the following property of logconcave (logarithmic concave) functions:[20] Let g(t) and G(t) denote respectively the probability density function and the cumulative distribution function of a random variable, if g(t) is logconcave in t, then the hazard function $g(t)/[1-G(t)]$ is an increasing function of t.[21] To apply this result, let $g(t) = at^b e^{-ct}$. Clearly $b > 0$ implies that g(t) is logconcave in t since $\partial^2\{\log[g(t)]\}/\partial t^2 = -b/t^2 < 0$. Thus, the hazard rate

[20] A function y(x) defined on a convex set X is logconcave if for any $x_1, x_2 \in X$ and any $\lambda \in [0,1]$, $y(\lambda x_1+(1-\lambda)x_2) \geq [y(x_1)]^{\lambda}[y(x_2)]^{(1-\lambda)}$. If $y(x) \geq 0$, then y(x) is logconcave if $\log[y(x)]$ is concave.

[21] This can be proved by using Prékopa's theorem (Prékopa, 1971, Theorem 2) to show that the logconcavity of g(t) implies the logconcavity of $[1-G(t)]$, and the fact that $d\{g(t)/[1-G(t)]\}/dt = -d^2\{\log[1-G(t)]\}/dt^2$.

of arrest corresponding to the fitted age-crime profile g(t) is increasing in t (i.e., positive duration dependence), which is consistent with the implications of the model.[22]

The third set of evidence is drawn from the age-crime profile itself. The age-crime profile describes how the probability density function f(t) varies with t. What is the shape of the hazard function corresponding to the pdf f(t)? To obtain the hazard function from f(t), one needs to find the ratio f(t)/[1-F(t)]. It is possible to get a rough estimate of this ratio from aggregate crime statistics. Figures 6 through 8 depict the hazard rates of arrest calculated from the data in Quetelet (1984). Figure 6 corresponds to the age-crime profile in Figure 1; and Figures 7 and 8 correspond respectively to Figures 2 and 3. Details on the derivations of these curves are described in appendix A. All the five curves show that the hazard rates of arrest are in general increasing in age, which again support the implications of the model.

In sum, the findings are consistent with the implications of the model. The first set of evidence supports the implication that the intensity rate of offending increases with time, while the last two sets of evidence support the implication that the hazard rate of arrest increases with time. Clearly, these findings do not support the prevailing interpretation. Hence, the empirical evidence strengthens the credibility of the model as a description of actual dynamic criminal behavior.

---

22 Theoretically, it is obvious that probability density functions of the form $g(t) = at^b e^{-ct}$ must have $b > 0$ in order to fit the age-crime profile. This is because $g'(t) = [b-ct]at^{b-1}e^{-ct}$, so that $b > 0$ is a necessary condition for fitting the rising part of the age-crime profile. A positive b implies the logconcavity of g(t), hence the hazard rate increases with t for both the gamma distribution and the three-parameter extensions.

V. Conclusion

In the past two decades, a great deal of criminological research has been focused on the longitudinal aspects of criminal behavior.[23] Economists, however, have paid little attention to these longitudinal problems. The present study provides a first systematic economic investigation of one important longitudinal aspect of criminal behavior: the age-crime profile. The analysis demonstrates that criminal decision can be treated as a dynamic consumption-investment problem. It substantiates Winship and Rosen's (1988) observation that economic theory can contribute to the literature by providing fundamental insights for the problem.

The paper begins with a brief review of the massive literature on the age-crime profile first discovered by Quetelet more than 150 years ago. A dynamic model is formulated to offer an economic interpretation of the age-crime profile. Since the time to arrest is a duration variable and an offender can continue participating in criminal activity only if he has not yet been arrested, the hazard function approach is used to model the criminal decision problem. It is crucial to model the decision problem through the hazard function (a conditional probability) because criminal decision is made at each point in time, taking the past as given, and is therefore conditional in nature. When some partial information is available (the offender has not yet been arrested), the desired probabilities are

---

[23] The emphasis on longitudinal issues in criminology is exemplified by the way federal research fundings are allocated (Gottfredson and Hirschi, 1986, 1988). The most notable current example is the massive 5 to 8 years overlapping longitudinal study, which costs more than 2 million dollars, jointly funded by the National Institute of Justice and the MacArthur Foundation. It is claimed that the study is one of the most complete longitudinal and interdisciplinary studies in the history of criminology (National Institute of Justice, 1990).

necessarily conditional ones.[24]

By focusing on the intensity aspect, the model shows that one cannot infer the time path of the intensity rate of offending directly from the age-crime profile. Contrary to the prevailing interpretation which argues that the falling part of the age-crime profile reflects a decline in the intensity rate of offending, the model shows that the intensity rate increases with time and does not decline at all. The age-crime profile is falling in a certain age range because of selection effect: fewer offenders are arrested in this range because a large percentage of offenders have already been arrested. An important result of the paper is that the model can generate the age-crime profile, which means that the age-crime profile can be generated by variations in the intensity rate alone. The implications of the model are also found to be consistent with empirical findings.

One of the goals of this paper is to introduce the age-crime profile into economists' research agenda. Many major issues remain to be addressed. The present analysis, being a first and exploratory endeavor, has its own limitations. Future work should consider at least three extensions: heterogeneity, participation, and recidivism. The paper assumes homogeneous offenders. Whether heterogeneity plays an important role in determining the shape of the age-crime profile remains to be explored. Paralleling to the analysis of the intensity aspect, it is useful to study the shape of the age-participation profile and its relation with the age-crime profile. The effects of recidivistic behavior on the intensity and the participation decisions should also be examined.

---

[24] See Heckman and Singer (1986), Kiefer (1988), or Pudney (1989) for a discussion of hazard function modeling. The importance of conditioning is emphasized in Ross (1989).

Appendix A: Derivations of Figures 1-3 and 6-8

Figures 1-3 and 6-8 are derived from the data reported in Tables 12 and 13 in Quetelet (1984). The age on the X-axis in each figure is the mid-point of the "person's age" (column 1 of Table 12). The variable(s) on the Y-axis for each figure is (are) obtained by the procedure indicated below.

Figure 1: The crime rate on the Y-axis is the "degrees of the propensity for crime" in column 6 of Table 12.

Figure 2: The crime rate for crime against persons on the Y-axis is obtained by dividing the "crimes against persons" (column 2 of Table 12) by the "population according to ages" (column 5 of Table 12). The crime rate for crime against property is obtained by dividing the "crimes against property" (column 3 of Table 12) by the "population according to ages" (column 5 of Table 12). Since the ranges of the two crime rates differ quite substantially, two separate scales for the crime rates are used.

Figure 3: The crime rate for men on the Y-axis is the "degrees of the propensity for crime" for "men" in column 6 of Table 13. The crime rate for women is the "degrees of the propensity for crime" for "women" in column 7 of Table 13.

Figure 6: Let f_i be the crime rate (column 6 of Table 12) for age group i ($i=1,2,\ldots,14$), then the standardized crime rate f_i^* is obtained by the formula $f_i^* = f_i/(\Sigma_{j=1}^{14} f_j)$. The purpose of the standardization is to make the standardized crime rates (f_i^*) sum up to one so that it is a well defined distribution function. Then the hazard rate of arrest on the Y-axis, h_i for age group i ($i=1,2,\ldots,13$), is obtained by the formula $h_i = f_i^*/[1-\Sigma_{j=1}^{i} f_j^*]$.

Figures 7-8: The procedure is similar to that of Figure 6 and the only change is the crime rate for each curve. The ranges of the hazard rates of arrest of men and women differ quite substantially, thus two separate scales are created for the hazard rates in Figure 8.

Appendix B: Simplifying the Objective Function

Let $D(c,t) = [e^{-rs(c)(T-t)} - e^{-r(T-t)}]/r$. To simplify (2), notice first that π_2 and π_3 do not depend on z, hence straightforward integration yields

$$\int_t^{t+s(T-t)} e^{-r(z-t)}\pi_2 dz = [1-e^{-rs(T-t)}]\pi_2/r, \tag{A1}$$

and $$\int_{t+s(T-t)}^T e^{-r(z-t)}\pi_3(c(t))dz = D(c(t),t)\pi_3(c(t))/r. \tag{A2}$$

Integration by parts and $F(0) = 0$ implies

$$\int_0^T e^{-rt}[1-F(t)]dt = -e^{-rT}[1-F(T)]/r + 1/r - \int_0^T [e^{-rt}f(t)/r]dt. \tag{A3}$$

In addition, $\int_0^T [e^{-rT}f(t)/r]dt = e^{-rT}F(T)/r$ implies

$$\int_0^T e^{-rt}D(c(t),t)f(t)dt = \int_0^T [e^{-rt}e^{-rs(T-t)}f(t)/r]dt - e^{-rT}F(T)/r. \tag{A4}$$

It follows from (A1), (A3), and (A4) that

$$\int_0^T e^{-rt}[\int_t^{t+s(T-t)} e^{-r(z-t)}\pi_2 dz]f(t)dt$$

$$= [1-e^{-rT}]\pi_2/r - \int_0^T e^{-rt}\pi_2\{[1-F(t)]+D(c(t),t)f(t)\}dt. \tag{A5}$$

Substitute (A2) and (A5) into (2), and replace f(t) by (1), (3) is obtained.

Appendix C: Proof of Proposition 1

(i) Solving the differential equation (7), $\lambda(t)[1-F(t)] - \lambda(0)[1-F(0)] = \int_0^t e^{-rx}U(c(x),x)[1-F(x)]dx$. Thus $\lambda(0) = -\int_0^T e^{-rx}U(c(x),x)[1-F(x)]dx = -\mathrm{Max}_{\tilde{c}}\ J(\tilde{c},F(0),0)$, by virtue of the assumption $F(0) = 0$ and the transversality condition $\lambda(T) = 0$. Hence, $\lambda(t)[1-F(t)] = -\int_t^T e^{-rx}U(c(x),x)[1-F(x)]dx$. It follows that

$$\lambda(t) = -\int_t^T e^{-rx}U(c(x),x)[(1-F(x))/(1-F(t))]dx$$

$$= -e^{-rt}\int_t^T e^{-r(x-t)}U(c(x),x)[(1-F(x))/(1-F(t))]dx$$

$$= -e^{-rt}[\mathrm{Max}_{\tilde{c}}\ J(\tilde{c},F(t),t)]$$

(ii) Let $J(0,F(t),t)$ denote the objective function when $c(x) = 0$ for $x \geq t$. By assuming an interior solution, $\mathrm{Max}_{\tilde{c}}\ J(\tilde{c},F(t),t) > J(0,F(t),t) = \int_t^T e^{-r(x-t)}\pi_3(0)dx = [1-e^{-r(T-t)}]\pi_3(0)/r \geq D(c(t),t)\pi_3(c(t))$, since $\pi_1(0) = \pi_3(0)$, $[1-e^{-r(T-t)}]/r \geq D(c(t),t)$, and $\pi_3' \leq 0$. It follows from (10) that $\lambda(t) + e^{-rt}D(c(t),t)\pi_3(c(t)) < 0$, for $t < T$.

Appendix D: General Case: s and π_3 depend on c

In this case, it is easy to verify that $c(t) \leq c(T)$ still holds for $t \leq T$. Differentiate (9) with respect to t, and replace λ' and λ by (7) and (9) respectively,

$$\tilde{M}[c'(t)] = e^{-rt}[(r+h)U_1 - U_{12} - Uh'] = e^{-rt}\tilde{Z}(c,t), \tag{A6}$$

where

$$\begin{aligned}\tilde{M} = e^{-rt}[&\pi_1'' - \theta''h - 2\theta'h' - \theta h'' \\ &+ D_{11}\pi_3 h + D\pi_3''h + D\pi_3 h'' + 2D_1\pi_3'h + 2D\pi_3'h' + 2D_1\pi_3 h'] + \lambda h'',\end{aligned} \tag{A7}$$

with $D_1 = -s'(T-t)e^{-rs(T-t)}$, $D_{11} = [-s''+r(s')^2(T-t)](T-t)e^{-rs(T-t)}$; and $\tilde{Z}(c,t) = W_1 + W_2$, where

$$W_1 = (r+h)(\pi_1'-\theta'h-\theta h') - (\pi_1-\theta h)h', \text{ and}$$

$$\begin{aligned}W_2 = &(1-s)e^{-rs(T-t)}\pi_3 h' - s'[1+(r+h-rs)(T-t)]e^{-rs(T-t)}\pi_3 h \\ &+ [Dh-(1-s)e^{-rs(T-t)}]\pi_3'h.\end{aligned}$$

Proposition 3: If (i) $D\pi_3$ is concave in c and (ii) $(r+h)\pi_1'-\pi_1 h' < 0$, then $c'(t) \geq 0$ for $t \in [0,T]$.

Proof: First, rearrange the right-side of (A7),

$$\begin{aligned}\tilde{M} = e^{-rt}[&\pi_1'' - \theta''h - 2\theta'h' - \theta h'' \\ &+ (D_{11}\pi_3 + 2D_1\pi_3' + D\pi_3'')h + 2D\pi_3'h' + 2D_1\pi_3 h'] + (\lambda + e^{-rt}D\pi_3)h''.\end{aligned}$$

Since $D\pi_3$ is concave in c (condition (i)) and $D_1 \leq 0$, thus $\tilde{M} < 0$ and the second-order sufficient condition is satisfied. Suppose $c'(t) < 0$ for some t $\in [0,T]$, then c(t) must attain a local minimum at some $t^* \in (0,T)$, because $c(t) \leq c(T)$ for all $t \in [0,T]$. This implies $c'(t^*) = 0$ and $c''(t^*) \geq 0$. Totally differentiate (A6) with respect to t,

$$(\partial\tilde{M}/\partial c)[c'(t)]^2 + (\partial\tilde{M}/\partial t)[c'(t)] + \tilde{M}[c''(t)]$$

$$= -re^{-rt}\tilde{Z} + e^{-rt}(\partial\tilde{Z}/\partial c)[c'(t)] + e^{-rt}(\partial\tilde{Z}/\partial t)$$

Evaluate the above expression at $c(t^*)$ and t^*, and notice that $c'(t^*) = 0$ (which implies $\tilde{Z}(c,t) = 0$ at $t = t^*$),

$$\tilde{M}[c''(t)] = e^{-rt}(\partial\tilde{Z}/\partial t) \quad \text{at } t = t^*. \tag{A8}$$

Now, $\partial\tilde{Z}/\partial t = rs(1-s)e^{-rs(T-t)}\pi_3 h'$

$$- s'\{rs[1+(r+h-rs)(T-t)] - (r+h-rs)\}e^{-rs(T-t)}\pi_3 h$$

$$+ \{[se^{-rs(T-t)}-e^{-r(T-t)}]h - rs(1-s)e^{-rs(T-t)}\}\pi_3' h$$

$$= rsW_2 - (1-s)e^{-r(T-t)}\pi_3' h^2 + s'(r+h-rs)e^{-rs(T-t)}\pi_3 h \tag{A9}$$

Condition (ii) implies $W_1 < 0$. Since $\tilde{Z} = W_1 + W_2 = 0$ at $t = t^*$, thus $W_2 > 0$ at $t = t^*$. Since $s' \geq 0$ and $\pi_3' \leq 0$, it follows from (A9) that $\partial\tilde{Z}/\partial t > 0$ at $t = t^*$, and hence (A8) implies $c''(t^*) < 0$. Contradiction. Hence, $c'(t) \geq 0$.

Remark: In this general case, the intensity rate also increases with time. The presence of an endogenous length of incapacitation and an endogenous post-incapacitation returns considerably complicate the optimal control problem so that condition (i) of Proposition 3 is required to satisfy the second-order sufficient condition. It is easy to check that there are many other conditions which also guarantee $\tilde{M} < 0$. Condition (ii) is not a restrictive one, as there are many examples which satisfy it, e.g., $\pi_1(c) = r + (1+c)^{1/2}$ and $h(c) = Hc$ $(H > 1)$ (or $h(c) = c\exp[1/(1-c)]$). Alternatively, it can be replaced by the weaker condition $W_1 < 0$.

References

Blumstein, Alfred, and Cohen, Jacqueline. "Estimation of Individual Crime Rates from Arrest Records." Journal of Criminal Law and Criminology 70 (1979): 561-585.

Blumstein, Alfred; Cohen, Jacqueline; Das, Somnath; and Moitra Soumyo. "Specialization and Seriousness During Adult Criminal Careers." Journal of Quantitative Criminology 4 (1988): 303-345.

Blumstein, Alfred; Cohen, Jacqueline; and Farrington, David. "Criminal Career Research: Its Value for Criminology." Criminology 26 (1988a): 1-35.

Blumstein, Alfred; Cohen, Jacqueline; and Farrington, David. "Longitudinal and Criminal Career Research: Further Clarifications." Criminology 26 (1988b): 57-74.

Blumstein, Alfred; Cohen, Jacqueline; Roth, Jeffrey; and Visher, Christy. Criminal Careers and "Career Criminals". (2 volumes) Washington, D.C.: National Academy Press, 1986.

Blumstein, Alfred, and Graddy, Elizabeth. "Prevalence and Recidivism in Index Arrests: A Feedback Model." Law and Society Review 16 (1982): 265-290.

Christensen, Ronald. "Projected Percentage of U.S. Population with Criminal Arrest and Conviction Records." In Task Force Report: Science and Technology, edited by The President's Commission on Law Enforcement and Administration of Justice. Washington, D.C.: U.S. Government Printing Office, 1967.

Cohen, Jacqueline. "Research on Criminal Careers: Individual Frequency Rates and Offense Seriousness." In Criminal Careers and "Career Criminals", vol. 1, edited by Alfred Blumstein et al. Washington, D.C.: National Academy Press, 1986.

Davis, Michael. "Time and Punishment: An Intertemporal Model of Crime." Journal of Political Economy 96 (March 1988): 383-390.

DeFleur, Lois. Delinquency in Argentina. Pullman: Washington State University, 1970.

Farrington, David. "Offending from 10 to 25 Years of Age." In Prospective Studies of Crime and Delinquency, edited by K. Van Dusen and S. Mednick. Boston: Kluwer-Nijhoff, 1983.

Farrington, David. "Age and Crime." In Crime and Justice, vol. 7, edited by M. Tonry and N. Morris. Chicago: University of Chicago Press, 1986.

Fry, Lincoln. "Drug Abuse and Crime in a Swedish Birth Cohort." British Journal of Criminology 25 (1985): 46-59.

Goring, Charles. The English Convict. London: Her Majesty's Stationery Office, 1913.

Gottfredson, Michael, and Hirschi, Travis. "The True Value of Lambda Would Appear to be Zero: An Essay on the Career Criminals, Criminal Careers, Selective Incapacitation, Cohort Studies, and Related Topics." Criminology 24 (1986): 213-234.

Gottfredson, Michael, and Hirschi, Travis. "Science, Public Policy, and the Career Paradigm." Criminology 26 (1988): 37-55.

Greenberg, David. "Age and Crime." In Encyclopedia of Crime and Justice, edited by Sanford Kadish. New York: Free Press, 1983.

Greenberg, David. "Age, Crime and Social Explanation." American Journal of Sociology 91 (1985): 1-21.

Heckman, James, and Singer, Burton. "Econometric Analysis of Longitudinal Data." In Handbook of Econometrics, vol. III, edited by Zvi Griliches and Michael Intriligator. New York: Elsevier, 1986.

Hirschi, Travis, and Gottfredson, Michael. "Age and the Explanation of Crime." American Journal of Sociology 89 (1983): 552-584.

Hirschi, Travis, and Gottfredson, Michael. "Age and Crime, Logic and Scholarship: Comment on Greenberg." American Journal of Sociology 91 (1985): 22-27.

Kiefer, Nicholas. "Economic Duration Data and Hazard Functions." Journal of Economic Literature 26 (1988): 646-679.

Killingsworth, Mark. Labor Supply, New York, NY: Cambridge University Press, 1983.

Loeber, Rolf, and Snyder, Howard. "Rate of Offending in Juvenile Careers: Findings of Constancy and Change in Lambda." Criminology 28 (1990): 97-109.

Messinger, Sheldon, and Bittner, Egon, eds. Criminology Review Yearbook. Vol. 1. California: Sage, 1979.

National Institute of Justice. "Massive Study Will Trace Developmental Factors that Cause or Prevent Criminality." NIJ Reports 220 (May/June 1990): 2-3.

Neison, Francis. Contributions to Vital Statistics. London: Simpkin, Marshall, 1857.

Prakasa Rao, B.L.S. Nonparametric Functional Estimation. Orlando: Academic Press, 1983.

Prékopa, András. "Logarithmic Concave Measures with Application to Stochastic Programming." Acta Sci Math (Szeged) 32 (1971): 301-316.

Press, William; Flannery, Brian; Teukolsky, Saul; and Vetterling, William. Numerical Recipes. New York: Cambridge University Press, 1986.

Pudney, Stephen. Modelling Individual Choice. New York: Basil Blackwell, 1989.

Quetelet, Adolphe. Research on the Propensity for Crime at Different Ages. (First Edition 1831, English Translation 1984) Translated by Sylvester, Sawyer. Ohio: Anderson, 1984.

Rosen, Sherwin. "A Theory of Life Earnings." Journal of Political Economy 84 (1976): S45-S67.

Ross, Sheldon. Introduction to Probability Models. San Diego: Academic Press, 1989.

Steffensmeier, Darrell; Allan, Emilie; Harer, Miles; and Streifel, Cathy. "Age and the Distribution of Crime." American Journal of Sociology 94 (January 1989): 803-831.

Stigler, George. "The Optimum Enforcement of Laws." Journal of Political Economy 78 (May/June 1970): 526-536.

Stigler, Stephen. The History of Statistics. Cambridge: Belknap Press, 1986.

Tittle, Charles. "Two Empirical Regularities (Maybe) in Search of an Explanation: Commentary on the Age/Crime Debate." Criminology 26 (1988): 75-85.

Wadsworth, Michael. "Delinquency in a National Sample of Children." British Journal of Criminology 15 (1975): 167-174.

Winship, Christopher, and Rosen, Sherwin. "Introduction: Sociological and Economic Approaches to the Analysis of Social Structure." American Journal of Sociology 94 (1988): S1-S16.

Wolfgang, Marvin; Figlio, Robert; and Sellin, Thorsten. Delinquency in a Birth Cohort. Chicago: University of Chicago, 1972.

Wolfgang, Marvin. "Some New Findings from the Longitudinal Study of Crime." Australian Journal of Forensic Sciences 13 (1980): 12-29.

Yaari, Menahem. "On the Consumer's Lifetime Allocation Process." International Economic Review 5 (1964): 304-317.

FIGURE 1
Age-Crime Profile

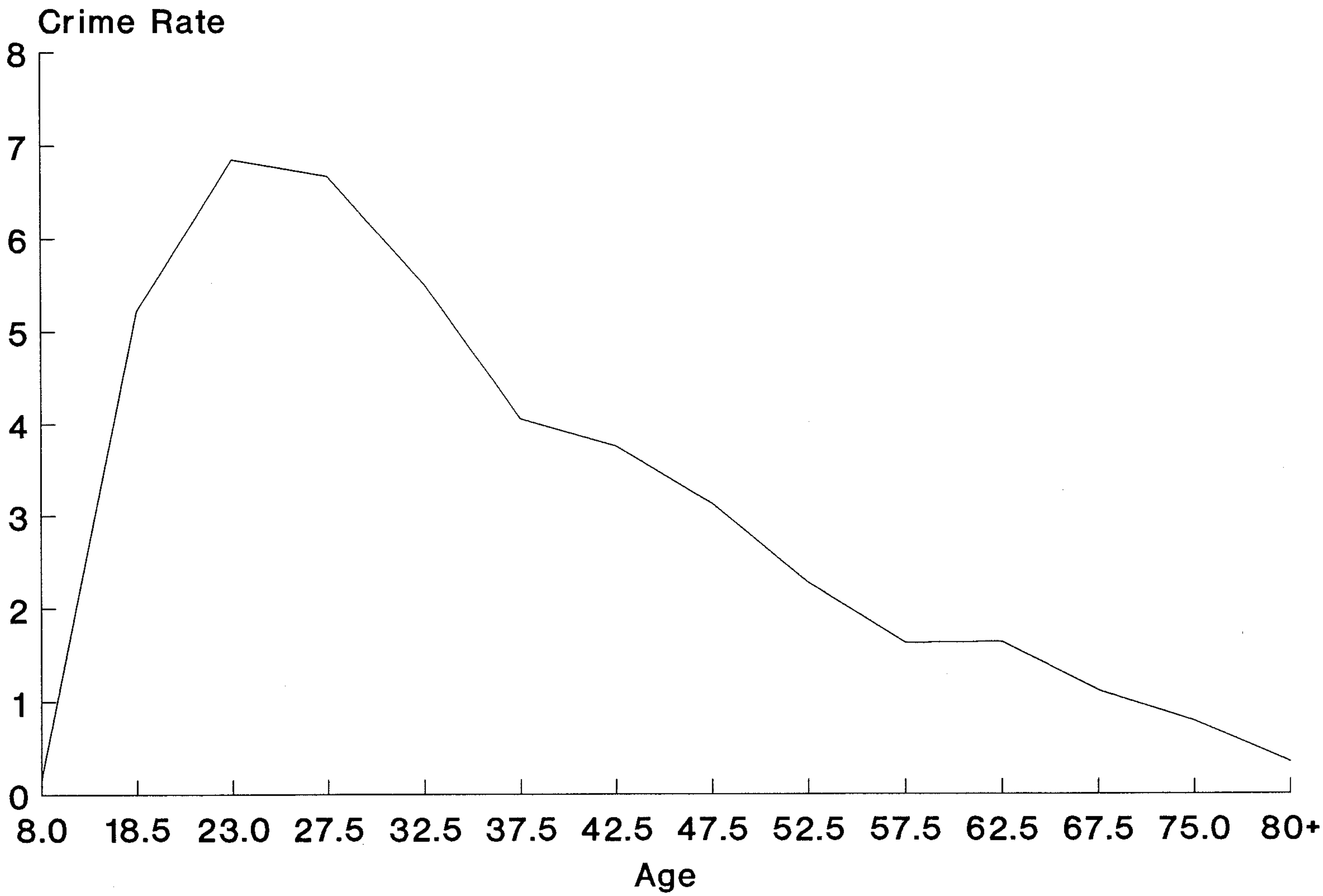

FIGURE 2
Age-Crime Profiles of Crimes
against Persons and Property

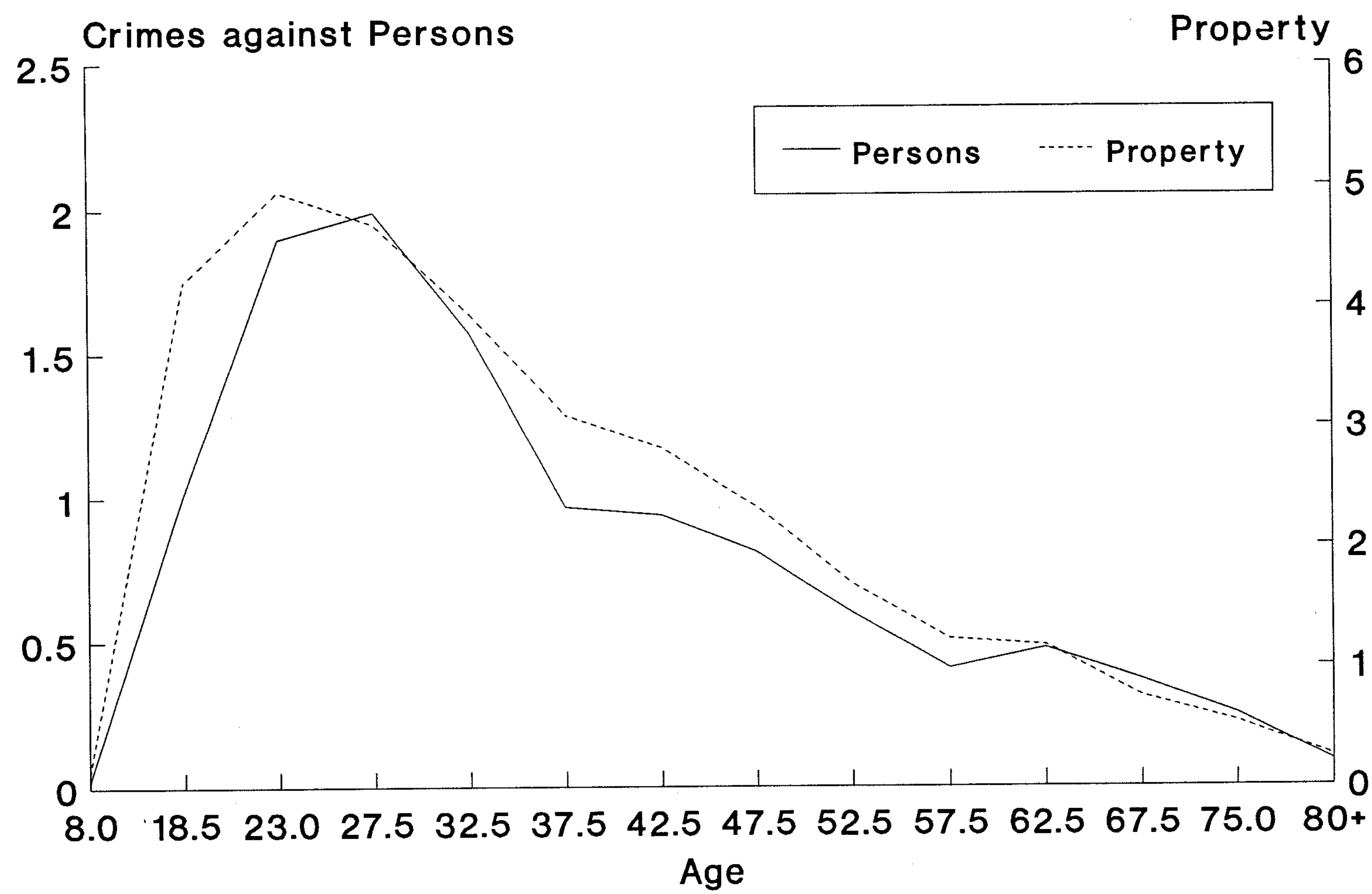

FIGURE 3
Age-Crime Profiles of Men and Women

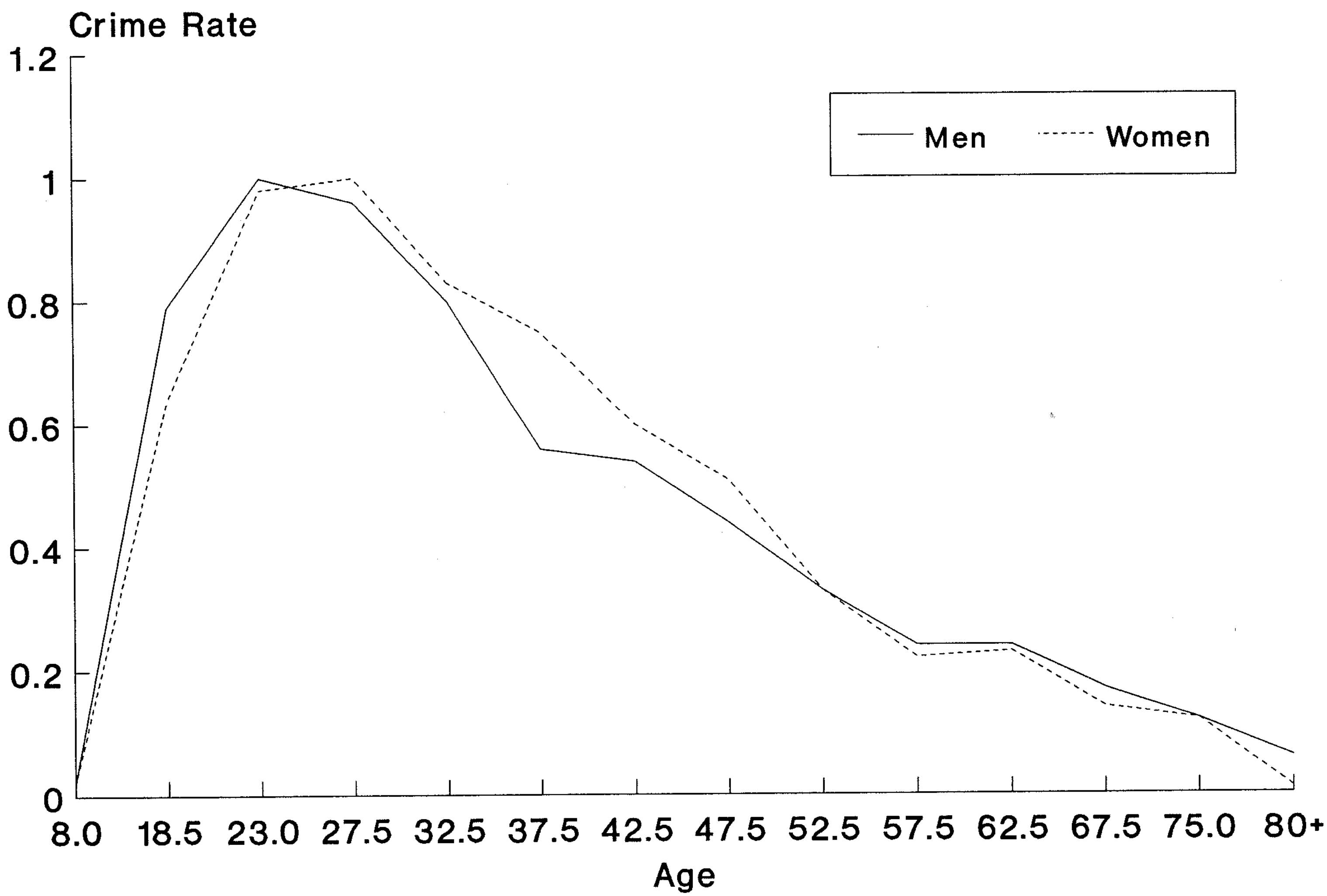

FIGURE 4
c(t) versus t

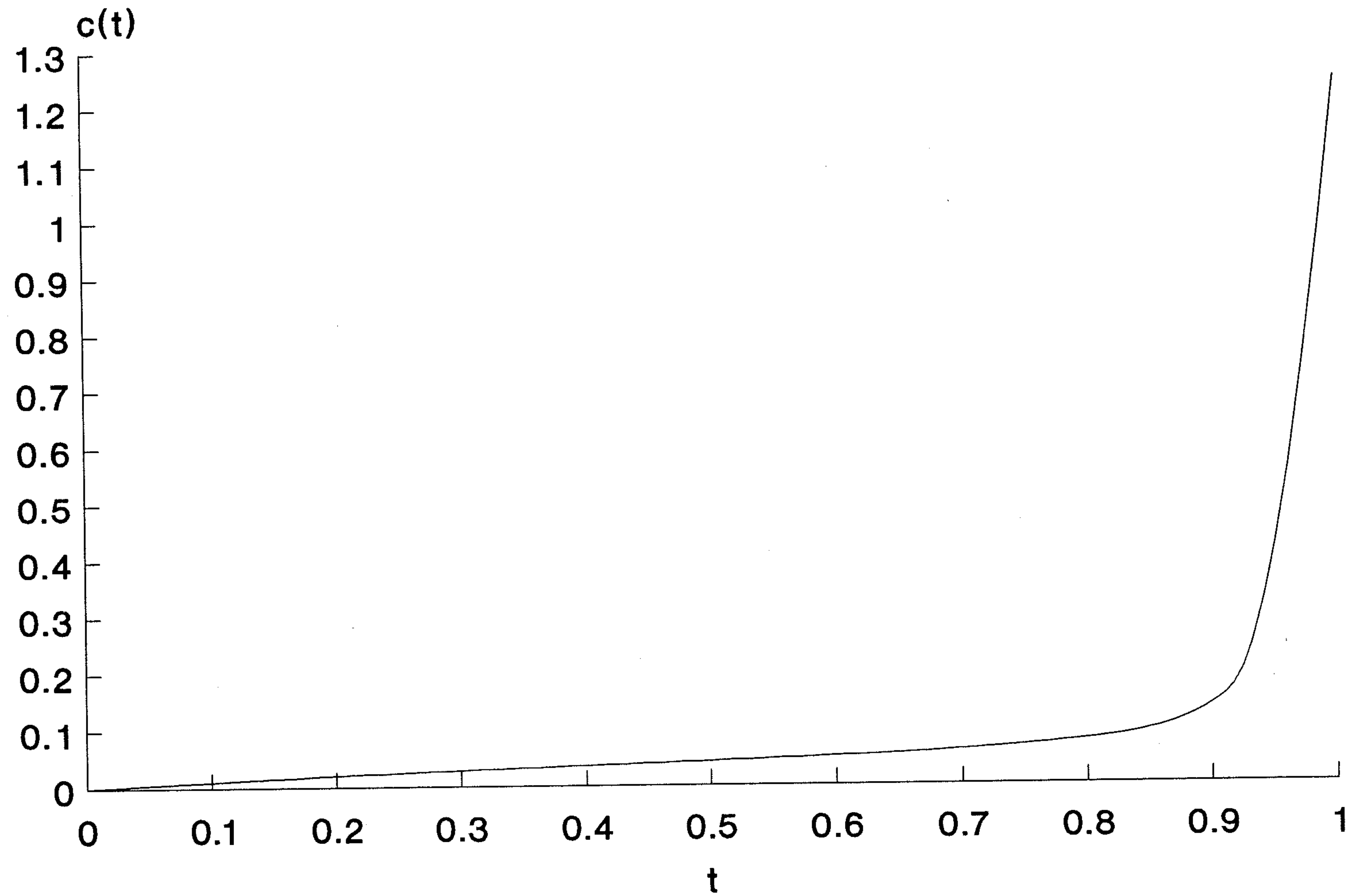

FIGURE 5
f(t) versus t

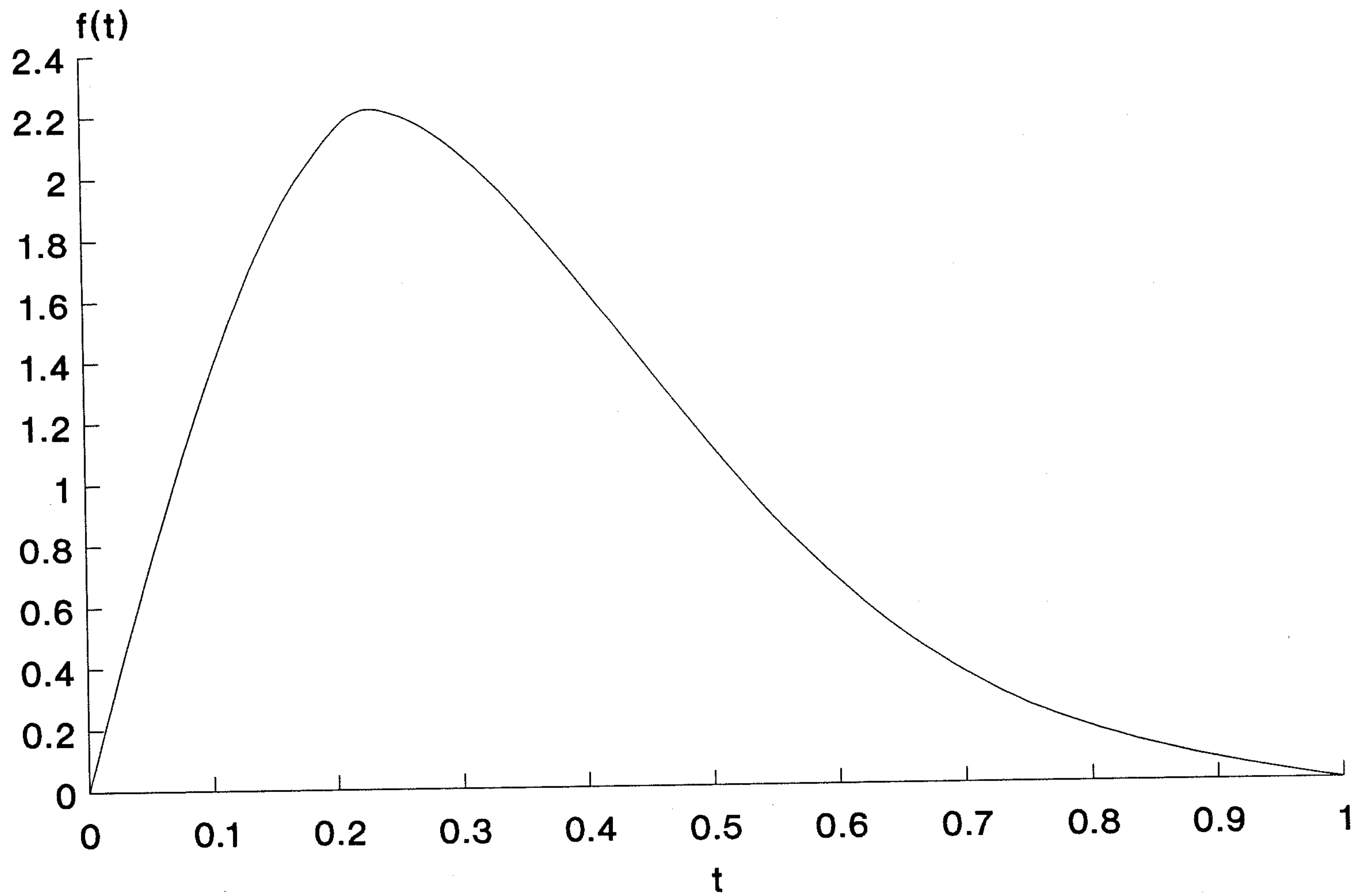

FIGURE 6
Age-Hazard Profile

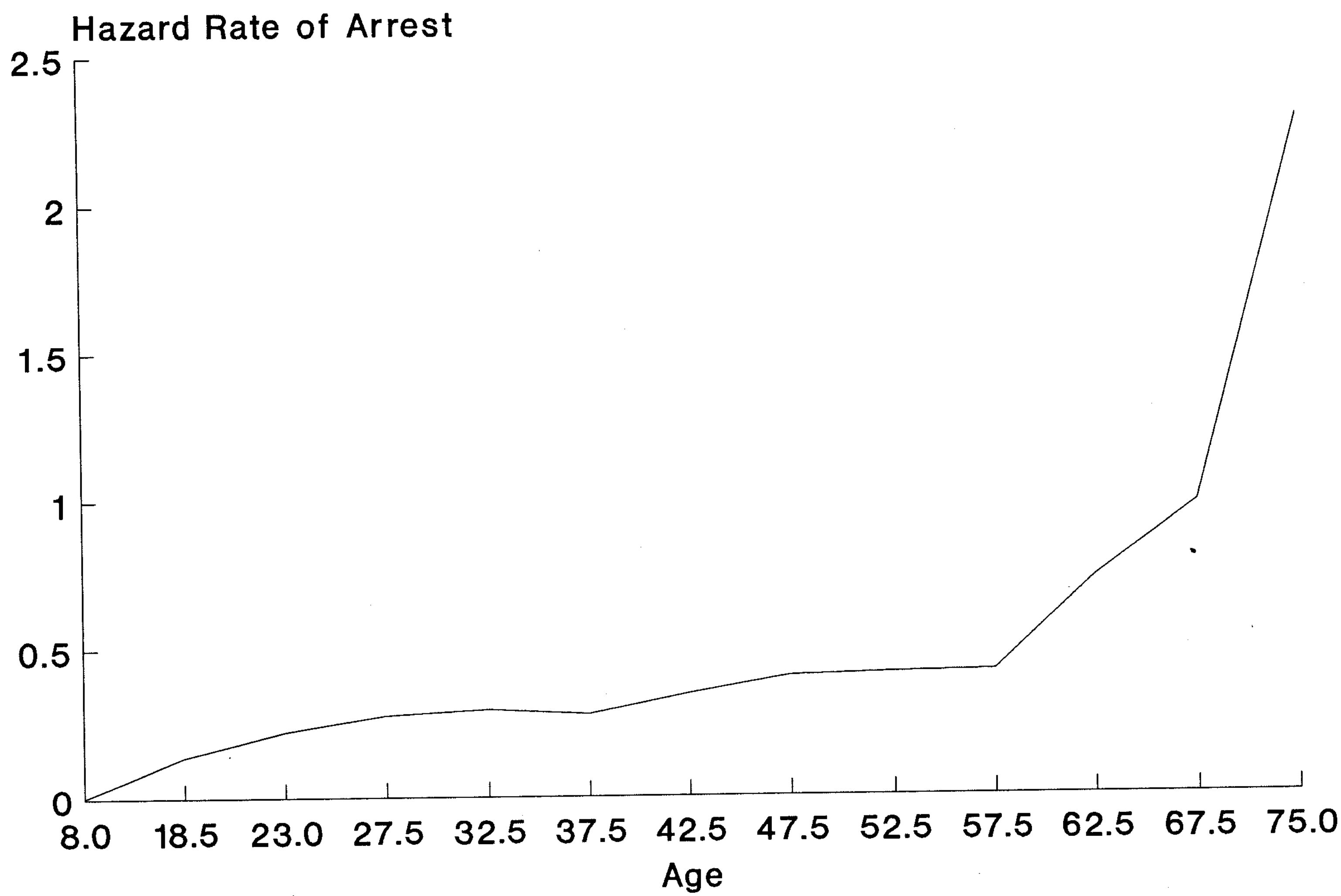

FIGURE 7

Age-Hazard Profiles of Crimes against Persons and Property

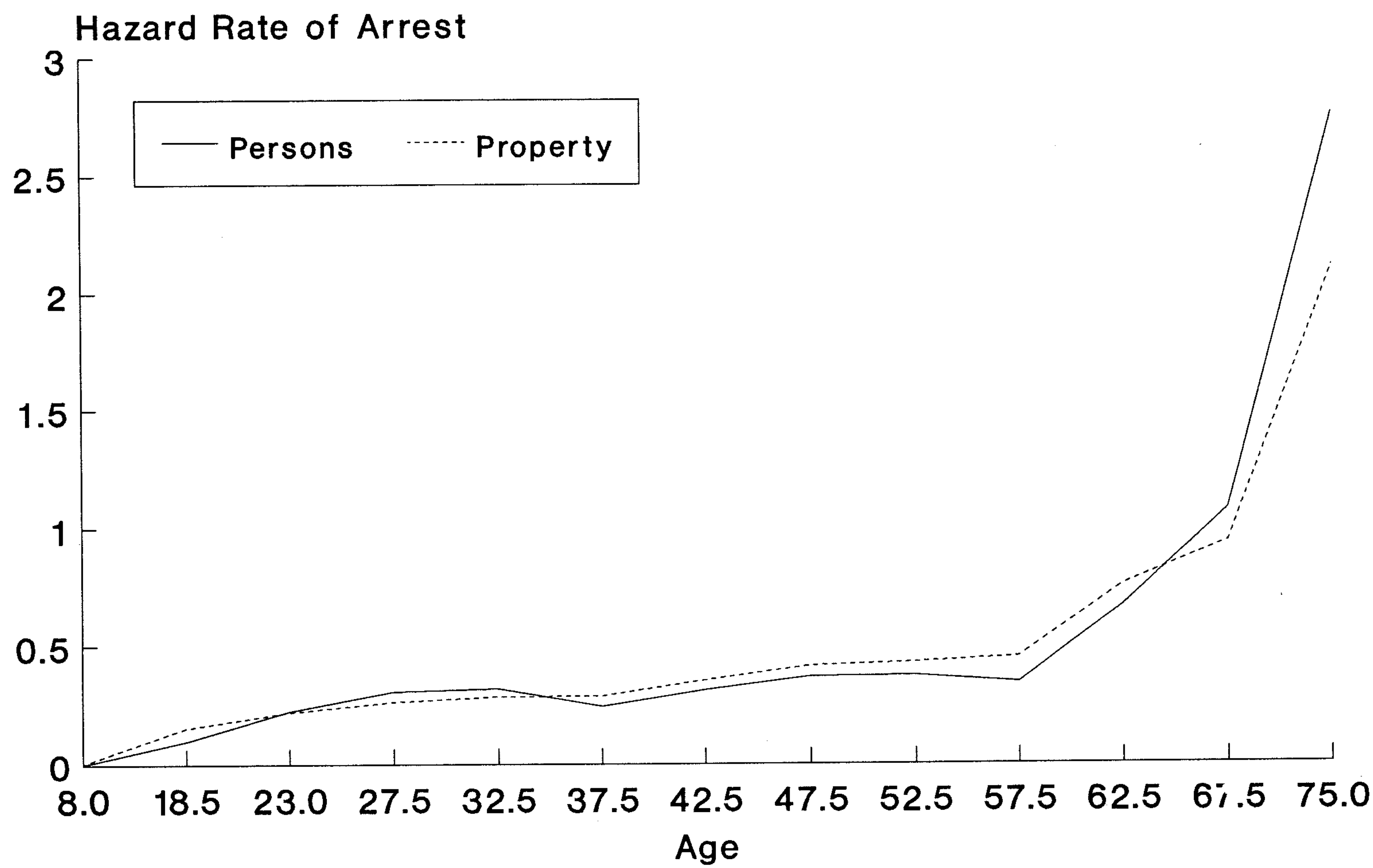

FIGURE 8
Age-Hazard Profiles of Men and Women