Coddling Fatalistic Criminals: A Dynamic Stochastic Analysis of Criminal Decision-Making

Brendan O'Flaherty, Columbia University

May 1997

Discussion Paper Series No. 9697-19

CODDLING FATALISTIC CRIMINALS:
A Dynamic Stochastic Analysis of Criminal Decision-Making

Brendan O'Flaherty
Department of Economics, Mail Code 3308
Columbia University
New York, NY 10027
May 1997

Abstract

Decisionmakers who confront a long sequence of criminal opportunities act differently from those who confront a single opportunity. If the sequence is long enough, people will take big chances in return for very small gains, even if the probability of detection is very great and the scale of punishment very large. Risk neutral people will appear to love risk. For long enough sequences of future opportunities, raising the probability of detection increases the amount of crime committed, rather than lowering it. Constitutional safeguards are an important deterrent to crime.

People are often presented with opportunities to commit crimes. In assessing each opportunity, they need to consider the past, the present, and the future -- the crimes they have already committed, the opportunity at hand, and the crimes they are likely to commit. Decision-making in this context is different from decision-making when a single criminal opportunity is being analyzed in isolation -- the sort of problem Becker (1968) studied. In this paper I show how it is different.

In particular, if the sequence of future opportunities is long enough, rational people will take big chances in return for very small gains, even if the probability of detection is very great and the scale of punishment very large. Risk neutral people will appear to be risk-loving. For long enough sequences of future opportunities, raising the probability of detection increases the amount of crime committed, rather than lowering it. For sequences of any length, increases in the probability that an innocent person will go free reduce crime by more than equivalent increases in the probability that a guilty person will be punished, and as the length of the sequence grows this difference grows. Constitutional safeguards -- "coddling criminals," in today's political parlance -- are thus an important deterrent to crime. In the traditional one-shot model, convicting wrongly is just as bad for deterrent purposes as acquitting wrongly (see, for instance Schrag & Scotchmer 1994); but as the horizon over which potential criminals plan grows longer, convicting wrongly becomes ever worse in its consequences. Deontology and long-run consequentialism agree.

The intuition behind these results lies in the notion of fatalism. Kremer (1996) and Mahal and O'Flaherty (1997) discuss how fatalism affects decisions to engage in behavior that could cause AIDS. If an injecting drug user, for instance, believes that sometime in the future the temptation to use an infected needle will be so large that he will give in and contract AIDS, then he might as well start using dirty needles now, since he will be infected no matter what he does in the current period. Fatalism works in the opposite direction of fear; thus, for instance, if fear predominates, a partially effective vaccine encourages risky behavior while if fatalism predominates a partially effective vaccine discourages risky behavior.

The key problem in applying the results from AIDS research to the study of criminal behavior is that contracting AIDS is an all-or-nothing, once-in-a-lifetime event; punishment for crimes can be more finely graded and dispersed over one's lifetime. The timing question turns out not to be a serious problem: what we need to consider is lifetime punishment. The gradation question is more serious: fatalism is present only if lifetime punishment is a concave function of lifetime detected crimes. Since lifetime punishment is bounded from above, concavity is not an unreasonable assumption.

The next section of this paper sets out the simplest model of long-term criminal behavior, one where crimes are detected either in the same period in which they are committed, or never. This is essentially a straightforward extension of Mahal and O'Flaherty (1996). Section 2 derives a series of results about willingness to commit crimes and risk aversion. Section 3 derives the ergodic distribution of potential criminals by number of convictions. Section 4 compares the deterrence effects of increases in the probability of correct and incorrect detection and shows the greater effectiveness of constitutional safeguards. Section 5 shows how changes in punishment affect crime, and section 6 concludes.

All models in this paper have infinite horizons. Mahal and O'Flaherty (1997) show that in the study of AIDS finite horizon models are not much different from infinite horizon ones. I therefore have omitted finite horizon extensions from this paper.

1. THE SETTING

I consider an individual decision problem. A decision-maker (DM) lives for an infinite number of discrete periods. In each period he must decide whether to commit a crime. The (instantaneous) payoff from abstaining from crime is zero. The (instantaneous) payoff from committing a crime is a random variable b. This random variable (the attractiveness of that period's criminal opportunity) is drawn independently each period from the same distribution $F(\cdot)$. I assume that $F(\cdot)$ is atomless and its support is the positive half-line. The DM knows the current period's value of b before he decides whether to commit a crime, but he does not know future values.

At the end of each period, there is a constant hazard $q>0$ that the DM will die (or be removed from the sequence of temptations). The DM maximizes the expected undiscounted sum of payoffs. Note that $q=1$ corresponds to the one-shot case that Becker studied.

The DM is also liable for punishment in each period, based on his choice of activity. At the end of each period, an exogenous legal system either convicts him of a crime or does not. If he has committed a crime, the probability that he will be punished is $(1-\alpha)$; so α is the probability that a guilty party will escape punishment. If he has not committed a crime, the probability that he will be punished is β; so β is the probability that an innocent person will be punished. The parameters α and β both represent mistakes that the legal system can make. I assume

$$\alpha + \beta < 1;$$

criminal behavior increases the probability of criminal conviction. Policies that give police and prosecutors more leeway seek to reduce α; policies that reduce police powers and raise the hurdles for conviction seek to reduce β. Since it is probably hard to reduce α without increasing β or vice versa, I will be interested in the relative responsiveness of criminal activity to changes in these two parameters.

In this paper I assume that the probability of conviction each period depends only on the activities in that period. A crime that goes unpunished in the period in which it is committed goes unpunished forever. This assumption can be relaxed without serious changes to the results.

Punishment following a conviction depends on the number of previous convictions. Let P

$$P: R_+ \rightarrow R_+$$

denote the lifetime punishment function: a person convicted of x crimes in his lifetime is sentenced to a total punishment of $P(x)$. Thus if a person who has already been convicted of x crimes in his life is convicted of another one, the punishment he will receive is

$p(x)=P(x+1)-P(x)$

I assume throughout that $P(x)$ is a weakly increasing function of x and hence that $p(x)$ is always nonnegative and sometimes positive. I will also be concerned whether $P(\cdot)$ is bounded, convex, linear or concave. Notice that I have not restricted the domain of $P(\cdot)$ to the integers, even though we will never encounter a DM with a fractional number of convictions. This is for convenience; in fact I assume that $p(x)$ is twice differentiable.

Let $B(x)$ denote the set of crime benefits b that will induce a DM who is behaving optimally with x prior convictions to commit a crime.

Let $V(x|q,\alpha,\beta)$ denote the value function: $V(x|\cdot)$ is the expected value of current and future payoffs for a DM who has already been convicted of x crimes and acts optimally from now on. Then the fundamental recursion equation is

(1)

$$V(x|q,\alpha,\beta)=\max_{s}\{\int_{b\varepsilon s}\{b-(1-\alpha)p(x)+(1-q)[(1-\alpha)V(x+1)+\alpha V(x)]\}dF(b)+\int_{b\notin s}\{-\beta p(x)+(1-q)[\beta V(x+1)+(1-\beta)V(x)]\}dF(b)\}$$

The maximization in (1) is simple and obvious: a criminal opportunity b should be taken -- that is, $b\varepsilon B(x)$ -- if and only if the expression in curly brackets in the first integrand is greater than the expression in curly brackets in the second integrand,

$$b-(1-\alpha)p(x)+(1-q)[(1-\alpha)V(x+1)+\alpha V(x)]$$

$$\geq -\beta p(x)+(1-q)[\beta V(x+1)+(1-\beta)V(x)].$$

So $B(x)$ is an upper half-line with minimum $b^*(x|q,\alpha,\beta)$, where

(2) $\quad b^*(x|q,\alpha,\beta)=(1-\alpha-\beta)[p(x)+(1-q)[V(x)-V(x+1)]].$

If $b \geq b^*(x|q,\alpha,\beta)$ the DM commits the crime; otherwise he abstains. Note that if q=1, the DM follows the Becker one-shot rule,

(3) $\quad b^*(x|q,\alpha,\beta) = (1-\alpha-\beta)p(x)$

We can substitute (1) into (2) and simplify in order to derive a fundamental recursion equation for $b^*(\cdot|\cdot)$. Such a recursion equation is our basic interest, since the probability of criminal behavior

$$1 - F(b^*(x|q,\alpha,\beta))$$

is monotonically (and negatively) related to $b^*(\cdot|\cdot)$. To do so, define the following function:

$$G(z,y) = \int_y^\infty [\min(b,z) - y]dF(b) \qquad \text{if } z \geq y$$

$$= -\int_z^\infty [\min(b,y) - z]df(b) \qquad \text{if } z<y.$$

Note that $G(\cdot)$ is continuous and differentiable in both its arguments with

(4) $\quad \begin{cases} \frac{\partial G}{\partial z} = 1 - F(z) \\ \frac{\partial G}{\partial y} = -(1-F(y)) \end{cases}$

Then we can simplify and write the following proposition:

<u>Proposition 1</u>:

(a) The minimum criminal opportunity $b^*(x|q,\alpha,\beta)$ satisfies the following recursion equation:

(5) $\quad b^*(x|q,\alpha,\beta) = (1-\alpha-\beta)qp(x) + (1-q)\{(1-\beta)b^*(x) + \beta b^*(x+1)$

$$-(1-\alpha-\beta)G(b^*(x),b^*(x+1))\}.$$

(b) A unique solution to (5) exists and is continuous and differentiable in all its arguments.

Proofs are gathered in the appendix.

2. WILLINGNESS TO COMMIT CRIMES

Proposition 1 gives the fundamental recursion equation we can use for tracing out willingness to commit crimes. How that willingness varies with x depends on the curvature of the lifetime punishment function P. If P is concave, then additional crimes bring less punishment, and so willingness to commit crimes increases as the number of convictions increases. Moreover, the cut-off is always lower than the Becker one-shot cut-off: one gain from being convicted of a crime is that future crimes will be punished less harshly.

Conversely, if the lifetime punishment function is convex, these results are reversed. The cut-off $b^*(\cdot|\cdot)$ rises as convictions increase, and is always higher than the Becker one-shot cut-off. If the lifetime punishment function is linear, the cut-off is independent of the number of convictions and equals the Becker one-shot cut-off. Formally:

<u>Proposition 2</u>:

(a) If $P(\cdot)$ is concave, $b^*(x|q,\alpha,\beta)$ is a decreasing function of x, and always less than the Becker one-shot rule given by (3).

(b) If $P(\cdot)$ is convex, $b^*(x|q,\alpha,\beta)$ is an increasing function of x, and always more than the Becker one-shot rule.

(c) If $P(\cdot)$ is linear, $b^*(x|q,\alpha,\beta)$ is constant with respect to x, and the Becker one-shot rule holds.

Note that if punishment is bounded and concave, someone who has been convicted of many crimes will not be much deterred by threats of punishment:

<u>Corollary to proposition 2</u>:

If P is concave and bounded, then

$$\lim_{x\to\infty} b^*(x|q,\alpha,\beta) = 0$$

Moreover,

<u>Proposition 3</u>:

$$\lim_{q\to 0} b^*(x|q,\alpha,\beta) = 0$$

for all x,α,β if P is concave

Thus with convex and bounded punishment and a long enough expected sequence of trials, any temptation is big enough to succumb to, because eventually you are likely to be convicted enough times that this particular transgression won't matter.

What is the actual shape of punishment functions? Some provisions -- leniency for first offenders, for instance -- suggest some convex portions, while others -- "three strikes and you're in" -- suggest concavity. A stigma that attaches to anyone who has a record makes the punishment function concave (see Freeman [1992] for evidence on such a stigma). Since punishment is bounded from above, it is unlikely that punishment functions are convex or linear, at least in their entirety. Thus concavity and boundedness are probably the more interesting properties, especially for limit arguments.

3. THE DISTRIBUTION OF DMs BY NUMBER OF CONVICTIONS

These basic equations allow us to calculate the distribution of DMs by number of convictions.

Consider an arbitrary number of convictions $x>0$, and let $s(x)$ denote the number (or share) of DMs who have that many convictions. DMs leave this state in two ways: $qs(x)$ of them die each period, and $(1\text{-}q)c(x)s(x)$ get convicted of another crime, where

$$c(x) = (1-\alpha)(1-F(b^*(x)) + \beta F(b^*(x))$$

the unconditioned probability that a DM with x prior convictions will be convicted this period. On the other hand,

$$(1-q)c(x-1)s(x-1)$$

people who had $(x\text{-}1)$ convictions last period were convicted and so entered state x. The steady state requires that entries equal exits and so

$$qs(x) + (1-q)c(x)s(x) = (1-q)c(x-1)s(x-1)$$

or

$$\frac{s(x)}{s(x-1)} = \frac{(1-q)c(x-1)}{(1-q)c(x)+q}$$

If punishment is concave or linear $c(x-1) \le c(x)$ and so $s(x) < s(x-1)$. The more convictions, the fewer the number of DMs in the steady state. If punishment is convex, this relationship may still hold for some x, but no general statement is possible.

4. CHANGES IN CONVICTION PROBABILITIES

What happens when conviction probabilities change? I confine my attention to concave punishment functions; results for convex ones are generally opposite.

The easiest kind of change to consider is a reduction in the probability of wrongful conviction, β. Roughly speaking, this causes two effects, both in the same direction. First, it means that the difference in expected convictions caused by abstaining on this particular trial,

$$1-\alpha-\beta$$

increases; so even if the punishment function were linear, there would be a greater gain from abstention (see Schrag and Scotchmer 1994). Second, it means that given any pattern of behavior on future trials, the number of expected lifetime convictions is less, and so the relevant region of the punishment function is steeper. This also increases the gain from abstention, and so both effects reduce the attractiveness of crime.

On the other hand, for reductions in α, the probability of wrongly escaping conviction, the two effects work in opposite directions. Changing α changes $(1-\alpha-\beta)$ the same way changing β does, and so the effect on convictions from abstaining in this period is the same. But reducing α increases, rather than reduces, the number of expected future convictions, and makes the relevant range of the punishment function a region where that function is flatter, rather than steeper. Thus whenever the punishment function is concave, a reduction in β reduces crime more than an equal reduction in α.

This difference in effects grows as the expected number of temptations grows. Indeed, if punishment is concave, reductions in α increase crime if the horizon is sufficiently long. The intuition is that the second effect becomes stronger than the first. Recall that with bounded punishment if your future number of expected convictions is large enough you might as well start committing crimes now since current crimes will make almost no difference to lifetime punishment; reducing α increases the strength of this fatalistic argument.

Let

$$A(x|q,\alpha,\beta)=\frac{\partial b^{*}(x|q,\alpha,\beta)}{\partial\alpha},$$

$$W(x|q,\alpha,\beta)=\frac{\partial b^{*}(x|q,\alpha,\beta)}{\partial\beta}$$

$$D(x|q,\alpha,\beta)=W(x|q,\alpha,\beta)-A(x|q,\alpha,\beta),$$

and say "fatalism predominates at (q, x)" whenever $A(x|q,\alpha,\beta)\geq 0$ (increasing the probability of convicting a guilty person increases crime). We can restate the above informal ideas as three propositions.

Proposition 4:

If P is concave, $W(x|q,\alpha,\beta)<0$ for all x,q,α,β

Proposition 5:

If P is concave, $D(x|q,\alpha,\beta) < 0$.

Proposition 6:

If P is concave, there exists a quit probability q^*, $1 > q^* > 0$, such that for all x and all $q \le q^*$, $A(x|q,\alpha,\beta) \ge 0$.

5. CHANGES IN THE PUNISHMENT FUNCTION

The other class of policies proposed to deter crime are those that make punishment "tougher." In the one-shot case, definition of tougher punishment is obvious and so is the argument for its efficacy in reducing crime. For the repeated case neither definition nor argument is so clear.

What is clear is that if $p(x)$ were to increase for all x, $b^*(x|\cdot)$ would, too; this follows fairly directly from (5). But if lifetime punishment is bounded, such an across-the-board increase is not feasible.

Consider first a change in only one value of $p(x)$. Implicitly such a change changes $P(x')$ for $x' > x$, but (5) shows that this change in $P(x')$ is irrelevant for behavior. Since $b^*(x+1|\cdot)$ depends only on $p(x)$ and on $b^*(x+k|\cdot)$ for higher values of k, $b^*(x+1|\cdot)$ does not change when $p(x)$ changes. Thus we can differentiate (5) with respect to $p(x)$ and rearrange to obtain

$$\frac{\partial b^*(x)}{\partial p(x)} = (1-\alpha-\beta)\cdot\frac{q}{q+(1-q)c(x)}$$

Note that in the one-shot Becker case

$$\frac{\partial p^*(x)}{\partial p(x)} = (1-\alpha-\beta),$$

and so since $\frac{q}{q+(1-q)c(x)} < 1$, one-shot punishment impacts are larger than punishment impacts in the repeated decision problem. Indeed, as the horizon grows (that is, $q \to 0$), the impact of changes in punishment vanishes.

Unless x=0, however, the impact of a change in punishment at x is not limited to DM in state x; it affects all DMs with fewer than x convictions as well. From (5),

$$\frac{\partial b^*(x-1)}{\partial b^*(x)} = (1-q)c(x) < 1$$

and so

$$\frac{\partial b^*(x-1)}{\partial p(x)} = (1-q)c(x)(1-\alpha-\beta)\frac{q}{q+(1-q)c(x)}$$

and in general

$$\frac{\partial b^*(x-k)}{\partial p(x)} = \frac{(1-\alpha-\beta)q}{q+(1-q)c(x)}(1-q)^k \prod_{i=0}^{k-1} c(x-i).$$

for k=1,. . . Thus the effect of a change in $p(x)$ is felt by all DMs with x or fewer convictions, but the fewer the number of convictions the smaller the impact. All these impacts are in the intuitive direction: punishment deters crime. (There is a curious dichotomy here. If $P(x')$ changes -- that is, if $x' > x$ -- then behavioral incentives don't change. If $P(x')$ stays the same -- that is, if $x' < x$ -- then behavioral incentives change.)

If there is a binding physical bound on maximum punishment, then a simple increase in one $p(x)$ is impossible. Consider therefore a compound perturbation of the punishment function; let $p(x_0)$ increase by δ and $p(x_1)$ decrease by δ. Changes like this are probably the kind observed most often. The impact depends on whether $x_0 > x_1$ or not.

Suppose $x_0 > x_1$; punishment increases after more crimes but decreases after fewer. For x between x_0 and x_1, only the increase at x_0 matters; thus $b^*(x')$ increases and crime decreases. But the effect wears off as x decreases. For $x' \leq x_1$, both effects matter, but the effect of the increase is more distant and so is less strong. So $b^*(x')$ decreases in this range and crime increases. So the overall effect on crime is ambiguous and depends on the distribution of DMs by number of convictions.

Suppose $x_0 < x_1$; punishment decreases after more crimes, but increases after fewer. An example would be "three strikes and you're in." The pattern is the opposite. For x between x_0 and x_1, crime increases; for $x' \leq x_0$, crime decreases. Once again the overall effect is ambiguous and depends on the conviction distribution.

If punishment is concave or linear, the steady state distribution of DMs by number of convictions slopes down. For $x_0 > x_1$, the crime cut-off $b^*(x)$ increases for states x that are less populated than the states at which the cut-off decreases. If increases and decreases were of the same size, and $F(\cdot)$ were uniform in the relevant range, crime would unambiguously increase. But neither of these conditions holds, and so even with the steady state assumption the overall impact is ambiguous. For $x_0 < x_1$, the crime decreases occur in more populated states than the crime increases, but the magnitudes of the changes in $b^*(x)$ work in the opposite direction.

Notice that both the effects of changes in α and changes in $P(\cdot)$ are ambiguous even though DMs are risk neutral. From (3) it is clear that in the Becker one-shot model if $\beta = 0$, all that matters for risk neutral DMs is the product $(1-\alpha)p(x)$. Thus empirical findings of greater elasticity of crime with respect to $(1-\alpha)$ than with respect to $p(x)$ have been taken as evidence that criminals, at least on the margin, are risk-loving. (See Becker 1968; Ehrlich 1975, 1977; Wolpin 1978)

In a dynamic context, however, these empirical findings are not inconsistent with risk-neutral criminals (or even risk-averse ones).

6. CONCLUSION

Repeated decision problems are qualitatively different from one-shot decision problems. Thus empirical work based on the one-shot model runs a high risk of being misleading. Longitudinal analysis is the natural and least dangerous way to study the incentives for criminal behavior.

Proof of Proposition 1:

(a) This part is essentially algebra. To save notation, I will omit the parameters (q,α,β) in $b^*(\cdot)$ and $V(\cdot)$. From (1),

$$V(x) = \int_{b^*(x)}^{\infty}\{b-(1-\alpha)p(x)+(1-q)(1-\alpha)V(x+1)+(1-q)\alpha V(x)\}dF(b)$$

$$+\int_{0}^{b^*(x)}(-\beta p(x)+(1-q)\beta V(x+1)+(1-q)V(x)\}dF(b)$$

$$=(1-q)V(x)+\int_{b^*(x)}^{\infty}\{b-(1-\alpha)p(x)+(1-q)(1-\alpha)[V(x+1)-V(x)]\}dF(b)$$

$$+\int_{0}^{b^*(x)}\{-\beta p(x)+(1-q)\beta[V(x+1)-V(x)]\}dF(b).$$

But from (2)

$$(1-q)[V(x+1)-V(x)]=\frac{b^*(x)}{1-\alpha-\beta}-p(x);$$

and so

$$V(x)=\int_{b^*(x)}^{\infty}\{b-(1-\alpha)\frac{b^*(x)}{1-\alpha-\beta}\}dF(b)+(1-q)V(x)$$

$$-\int_{o}^{b^*(x)}\beta\frac{b^*(x)}{1-\alpha-\beta}dF(b)$$

$$=(1-q)V(x)+\int_{b^*(x)}^{\infty}\{b-(1-\alpha)\frac{b^*(x)}{1-\alpha-\beta}\}dF(b)-(1-\int_{b^*(x)}^{\infty}.dF(b))\beta\frac{b^*(x)}{1-\alpha-\beta}$$

$$=(1-q)V(x)+\int_{b^*(x)}^{\infty}\{b-(1-\alpha)\frac{b^*(x)}{1-\alpha-\beta}\}dF(b)+\int_{b^*(x)}^{\infty}\frac{\beta b^*(x)}{1-\alpha-\beta}-\beta\frac{b^*(x)}{1-\alpha-\beta}$$

$$= (1-q)V(x) + \int_{b^*(x)}^{\infty} (b - b^*(x))dF(b) - \beta \frac{b^*(x)}{1-\alpha-\beta}.$$

Then

$$V(x) - V(x+1) = (1-q)[V(x) - V(x+1)] + \int_{b^*} [b - b^*(x)]dF(b) - \int_{b^*(x+1)} [b - b^*(x+1)]dF(b)$$

$$-\frac{\beta}{1-\alpha-\beta}[b^*(x) - b^*(x+1)]$$

$$= -p(x) + \frac{b^*(x)}{1-\alpha-\beta} - G(b^*(x), b^*(x+1)) - \beta\frac{b^*(x)}{1-\alpha-\beta} + \beta\frac{b^*(x+1)}{1-\alpha-\beta}$$

Substituting this expression in (2) yields (5).

(b) To prove existence, uniqueness, and differentiability, it is easiest to prove these properties for the value function $V(\cdot|\cdot)$. Existence, uniqueness, and differentiability for $b^*(\cdot|\cdot)$ then follow from (2).

Consider the transform $U_{q\alpha\beta}$ defined by

$$U_{q\alpha\beta}u(x) = \max_{s} \{ \int_{b \in s} \{b - (1-\alpha)p(x) + (1-q)[(1-\alpha)u(x+1) + \alpha u(x)]\}dF(b)$$

$$+ \int_{b \notin s} \{-\beta p(x) + (1-q)[\beta u(x+1) + (1-\beta)u(x)]\}dF(b)\}.$$

To save notation in this section, I will supress the subscripts on U. It is clear that if u is a continuous and differentiable function of x, q, α, β then Uu is a continuous and differentiable function x, q, α, β also. Hence all that remains to be shown is that U is a contraction mapping.

Use the sup norm:

$$\|u_1, u_2\| = \sup_x |u_1(x) - u_2(x)|.$$

Consider $Uu_1(\cdot)$ and $Uu_2(\cdot)$ and suppose w.l.o.g. that

$$\|Uu_1, Uu_2\| = Uu_1(x) - Uu_2(x) > 0.$$

But

$$Uu_1(x) - Uu_2(x) \le \int_{b\varepsilon s_2} \{b - (1-\alpha)p(x) + (1-q)[(1-\alpha)u_1(x+1) + \alpha u_1(x)]\} dF(b)$$

$$+ \int_{b\notin s_2} \{-\beta p(x) + (1-q)[\beta u_1(x+1) + (1-\beta)u_1(x)]\} dF(b)$$

$$- \int_{b\varepsilon s_2} \{b - (1-\alpha)p(x) + (1-q)[(1-\alpha)u_2(x+1) + \alpha u_2(x)]\} dF(b)$$

$$- \int_{b\notin s_2} \{-\beta p(x) + (1-q)[\beta u_2(x+1) + (1-\beta)u_2(x)]\} dF(b)$$

where s_2 denotes the set that maximizes value with u_2. The inequality follows because $Uu_1(x)$ is at least as great as value constrained to be in s_2. The right-hand side of this inequality simplifies to

$$Uu_1(x) - Uu_2(x) \le (1-q)\{[u_1(x+1) - u_2(x+1)][(1-\alpha)\int_{s_2} dF(b) + \beta(1 - \int_{s_2} dF(b)]$$

$$+[u_1(x) - u_2(x)][\alpha \int_{s_2} dF(b) + (1-\beta)(1 - \int_{s_2} dF(b)]\}$$

where the expression in curly brackets is just a weighted average of

$$[u_1(x+1) - u_2(x+1)]$$

and

$$[u_1(x) - u_2(x)]$$

and $(1-q) < 1$. Thus

$$||Uu_1, Uu_2|| < \max[|u_1(x) - u_2(x)|, |u_1(x+1) - u_2(x+1)|] \le ||u_1, u_2||$$

and so U is a contraction mapping. QED.

Proof of Proposition 2:

Consider the transform $T_{q\alpha\beta}$ defined by

(6)

$$T_{q\alpha\beta}u(x) = (1-\alpha-\beta)qp(x) + (1-q)\{(1-\beta)u(x) + \beta u(x+1) - (1-\alpha-\beta)G(u(x), u(x+1))\}.$$

To save notation in this section, I will supress the subscripts on T. From (3) and proposition 1(b), $b^*(\cdot|\cdot)$ is the unique fixed point of T. Thus to show that $b^*(\cdot|\cdot)$ is decreasing in x, it suffices to show that if $u(\cdot)$ is decreasing, so is $Tu(\cdot)$. The smae for increasing and constant. I will confine the proof to part (a). Proof of parts (b) and (c) is so similar as to be obvious.

Suppose then that P is concave and $u(x)$ is decreasing. Then

$$\frac{\partial Tu(x)}{\partial x} = (1-\alpha-\beta)qp'(x) + (1-q)\{\frac{\partial u(x)}{\partial x}(1-c_u(x)) + \frac{\partial u(x+1)}{\partial x}c_u(x+1)\}$$

where

$$c_u(x) = \beta F(u(x)) + (1-\alpha)(1-F(u(x))) > 0$$

is the unconditional probability of being convicted this period if you have x convictions and will accept a criminal opportunity b if and only if $b \ge u(x)$. Since $p'(x)$ is negative from concavity and both derivatives in curly brackets are negative by hypothesis

$$\frac{\partial Tu(x)}{\partial x} < 0.$$

Hence for all $x, q, \alpha, \beta, b^*(x|q,\alpha,\beta)$ is decreasing in x.

For the second claim, we use the following:

Lemma 1:

If P is concave, $V(x|q,\alpha,\beta)$ is an increasing function of x.

Proof of lemma:

Use the transform U from the proof of proposition 1, and suppose $u(x)$ is increasing in x.

$$\frac{\partial Uu(x)}{\partial x} = \int_{b\varepsilon s}[-(1-\alpha)p'(x)+(1-q)[(1-\alpha)u'(x+1)+\alpha u'(x)]]dF(b)$$

$$+\int_{b\notin s}[-\beta p'(x)+(1-q)[\beta u'(x+1)+(1-\beta)u'(x)]dF(b)>0$$

where we can ignore changes in the bounds of integration by the envelope theorem, and the inequality follows because all terms are positive. QED.

From the lemma, then

$V(x)-V(x+1)<0$

and so from (2)

$b^*(x|q,\alpha,\beta)=(1-\alpha-\beta)p(x)+(1-\alpha-\beta)(1-q)[V(x)-V(x+1)<(1-\alpha-\beta)p(x)$,

the Becker one-shot criterion. QED.

Proof of corollary to Proposition 2:

From proposition 2(a)

$b^*(x|q,\alpha,\beta)\le(1-\alpha-\beta)p(x)$

and so

$$\lim_{x\to\infty} b(x^{*}|q,\alpha,\beta) \leq \lim_{x\to\infty} (1-\alpha-\beta)p(x).$$

But if P is bounded

$$\lim_{x\to\infty} p(x) = 0$$

and so since $b^{*}(x|\cdot)$ is by construction nonnegative, the corollary follows. QED.

Proof of Proposition 3:

Consider the transform

$$T_{q\alpha\beta}u(x) = (1-\alpha-\beta)qp(x) + (1-q)\{(1-\beta)u(x) + \beta u(x+1) - (1-\alpha-\beta)G(u(x),u(x+1))$$

Let $q \to 0$

$$\lim_{q\to 0} T_{q\alpha\beta}u(x) = (1-\beta)u(x) + \beta u(x+1) - (1-\alpha-\beta)G(u(x),u(x+1)).$$

The fixed point for the limit transform is any constant function. Hence by continuity $b^{*}(x|q,\alpha,\beta)$ must be arbitrarily close to a constant function for q sufficiently close to zero. By the corollary to proposition 2, the only constant function that can behave appropriately for arbitrarily large x is

$b^{*}(x|q,\alpha,\beta) = 0$. QED.

Proof of Proposition 4:

From (6)

$$\frac{\partial T_{q\alpha\beta}u(x)}{\partial\beta} = -qp(x) + (1-q)[-u(x) + u(x+1) - G(u(x), u(x+1))]$$

which is always negative if $u(x)$ is a decreasing function. Consider the fixed point $b^*(\cdot|\cdot)$ of this transformation. From proposition 2 it is a decreasing function on the right-hand side of this equation; hence

$$\frac{\partial T_{q\alpha\beta}b^*(x|\cdot)}{\partial\beta} < 0.$$

But since $T_{q\alpha\beta}b^* = b^*$, the proposition follows. QED.

Proof of Proposition 5:

For any function $v(x,\alpha,\beta)$ define

$$\Delta(v)(x,\alpha,\beta) = \frac{\partial v}{\partial\beta} - \frac{\partial v}{\partial\alpha}.$$

Let $u(x)$ be a decreasing function. Then

$$\Delta(T_{q\alpha\beta}u)(x,\alpha,\beta) = (1-q)\{[-u(x) + u(x+1) - G(u(x), u(x+1)] + G(u(x), u(x+1)\}$$

$$= (1-q)[-u(x) + u(x+1)] < 0.$$

At the fixed point

$$\Delta(T_{q\alpha\beta}b^*)(x,\alpha,\beta) = D(x|q,\alpha,\beta) < 0. \text{ QED.}$$

Proof of Proposition 6:

Let $u(x)$ be a strictly decreasing function. Define

$$Q(u) = \min_x \frac{G(u(x), u(x+1))}{G(u(x), u(x+1)) + p(x)}$$

Since $u(x)$ is strictly decreasing, $Q(u)>0$ and clearly $Q(u)<1$ (since $p(x)>0$ for some x.

Since

$$\frac{\partial T_{q\alpha\beta} u(x)}{\partial \alpha} = -qp(x) + (1-q)G(u(x), u(x+1)),$$

$$\frac{\partial T_{q\alpha\beta}}{\partial \alpha} \geq 0 \text{ for all } q \leq Q(u).$$

Consider the mapping $T^*_{\alpha\beta} u = T_{Q(u)\alpha\beta}$. It is easy to see that this is a contraction mapping, since the proof of proposition 1 relied only on $q<1$, not any particular value of q. So $T^*_{\alpha\beta}$ has a fixed point $v_{\alpha\beta}$ and

$$0 < Q(v_{\alpha\beta}) = q^* < 1.$$

But then $b^*(x|q^*, \alpha, \beta) = v_{\alpha\beta}$ solves (6) and

$$\frac{\partial b^*(x|q, \alpha, \beta)}{\partial \alpha} \geq 0 \text{ for all } q \leq q^*. \text{ QED.}$$

REFERENCES

Becker, Gary S., 1968: "Crime and Punishment: An Economic Approach," Journal of Political Economy, 76, 169-217.

Ehrlich, Isaac, 1975: "The Deterrent Effect of Capital Punishment: A Question of Life and Death," American Economic Review, 65, 397-417.

----------, 1977: "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence," Journal of Political Economy, 85, 741-88.

Freeman, Richard B., 1992: "Crime and the Employment of Disadvantaged Youth," in George Peterson and Wayne Vroman, eds., Urban Labor Markets and Job Opportunity. Washington, D.C.: Urban Institute.

Kremer, Michael, 1996: "Integrating Behavioral Choice into Epidemiological Models of AIDS," Quarterly Journal of Economics, 111, 549-74.

Mahal, Ajay and Brendan O'Flaherty, 1997: "Fatalism," Columbia University discussion paper.

Schrag, J. and Suzanne Scotchmer, 1994: "Crime and Prejudice: The Use of Character Evidence in Criminal Trials," Journal of Law, Economics, and Organization, 10, 319-42.

Wolpin, Kenneth, 1978: "An Economic Analysis of Crime and Punishment in England and Wales," Journal of Political Economy, 86, 815-40.

1996-1997 Discussion Paper Series

Department of Economics
Columbia University
1022 International Affairs Bldg.
420 West 118th Street
New York, N.Y., 10027

The following papers are published in the 1996-97 Columbia University Discussion Paper series which runs from early November to October 31 of the following year (Academic Year).

Website: http://www.columbia.edu/dlc/wp/econ/index.html.

You may download any papers found on this site.

For Ordering Hardcopies:

Domestic orders for discussion papers are available for purchase at the cost of $8.00 (U.S.) Per paper and $140.00 (US) for the series.

Foreign orders cost $10.00 (US) per paper and $185.00 for the series.

To order discussion papers, please write to the Discussion Paper Coordinator at the above address along with a check for the appropriate amount, made payable to Department of Economics, Columbia University. Please be sure to include the series number of the requested paper when you place an order.

1996-97 Discussion Papers
Economics Department, Columbia University

Series no.	Title	Authors
9697-01	Fertility Behavior Under Income Uncertainty	Ranjan, P.
9697-02	Trade Restrictions, imperfect Competition and National Welfare with Foreign Capital Inflows	Ranjan, P.
9697-03	Restructuring an Industry during Transition: a Two-Priced Model	Ericson, R.
9697-04	A Conformity Test for Cointegration	Dhrymes, P.
9697-05	Low-Wage Employment Subsidies in a Labor-Turnover Model of the 'Natural Rate' (November 1996)	Hoon, H.T. Phelps, E.
9697-06	The Knowledge Revolution	Chichilnisky, G.
9697-07	The Role of Absolute Continuity in "Merging Opinions" and "Rational Learning"	Miller. R. Sanchirico, C.W.
9697-08	The Soviet Bloc and the Soviet Union: Why did they fall apart?	Desai, P.
9697-09	Regionalism and Multilateral Tariff Cooperation	Bagwell, K. Staiger, R.
9697-10	Supplier Relations and the Market Context: A theory of handshakes	McLaren, J.
9697-11	Vertical Foreclosure with the Choice of Input Specifications	Choi, J.P. Yi, S.
9697-12	Equilibrium Vertical Foreclosure with Investment	Choi, J.P. Yi, S.
9697-13	Going Global: Transition from Plan to Market in the World Economy	Desai, P.
9697-14	Reciprocity, Non-discrimination and Preferential Agreements in the Multilateral Trading System	Bagwell, K. Staiger, R.
9697-15	Russia's Transition Toward the World Economy: Is the Market Mechanism Working?	Desai, P.
9697-16	A Dynamic Analysis of Licensing: The "Boomerang" Effect and Grant-Back Clauses	Choi, J.P.

Series no.	Title	Authors
9697-17	An Economic Theory of GATT	Bagwell, K. Staiger R.W.
9697-18	Climate Consumption and Climate Pricing from 1940 to 1990	Cragg, M.I. Kahn, M.E.
9697-19	Coddling Fatalistic Criminals: A Dynamic Stochastic Analysis of Criminal Decision-Making	O'Flaherty, B.