Stigma and Self-Fulfilling Expectations of Criminality[*]

September 3, 1996

Eric Rasmusen

Indiana University School of Business, 1309 E. 10th Street, Bloomington, Indiana, 47405. Office: (812) 855-9219. Fax: 812-855-3354. Email: Erasmuse@.indiana.edu. For the latest versions of this and other papers, go to my Webpage, http://ezinfo.ucs.indiana.edu/∼erasmuse.

2000: Eric Rasmusen, Professor of Business Economics and Public Policy and Sanjay Subhedar Faculty Fellow, Indiana University, Kelley School of Business, BU 456, 1309 E 10th Street, Bloomington, Indiana, 47405-1701. Office: (812) 855-9219. Fax: 812-855-3354. Erasmuse@indiana.edu. Php.indiana.edu/∼erasmuse.

Abstract

A convicted criminal suffers not only from public penalties, but from stigma, the reluctance of others to interact with him economically and socially. Conviction can convey useful information about the convicted, which makes stigmatization an important and legitimate function of the criminal justice system quite apart from moral considerations. The magnitude of stigma depends on expectations and the crime rate, however, which can lead to multiple, pareto-ranked equilibria with different amounts of crime.

I. INTRODUCTION

The economic approach to crime accepts that internal motivations such as conscience are important, but focuses on more easily measured and manipulated external incentives such as criminal penalties. This approach is theoretically attractive, consistent with common sense, and has had some degree of success in explaining empirical variation in crime rates.[1] The pattern of crime in the United States seems at first glance to support the importance of punishment. Reported crimes rose by 294 percent from 1960 to 1980, while the number of new prisoners per crime fell by 58 percent. During the 1980's, however, the number of prisoners per crime had risen by 118 percent, yet crime still rose by 8 percent.[2]

The economic model of crime has been elaborated over the years, but we still do not have a satisfactory explanation for the decreased impact of criminal penalties. A number of articles have explored variants of what I will call the "overload theory": that when crime increases, law enforcement funding does not increase enough to prevent the expected penalty from declining, which increases crime still further.[3] The overload theory can explain how a society might move from an equilibrium with low crime to one with high crime, but the amount of crime is still mediated by the expected penalty, so it cannot explain the U. S. pattern.[4]

Over a period as long as thirty years, it may be that social influences are more important determinants of changes in crime rates than the incentives traditionally studied by economists. Some social influences, such as present-orientedness or conscience, are tastes, which economists take as given. Stigma is not. Stigma refers to someone's reluctance to interact with a someone else who has a criminal record. For the criminal, stigma is an external incentive, like a jail term, not an internal motivation, like conscience. Standard economic modelling can be used to ask how the criminal will respond to stigma and why people find it in their self-interest to treat criminals differently from noncriminals.

Stigma can be either economic (for example, a lower wage) or social (for example, difficulty finding a spouse). Economic stigma is the easier of these to measure, and a number of scholars have tried to estimate its impact.[5] Whether stigma is important seems to depend on the context. Lott, for example, finds a short-run income reduction of 39 percent from bank embezzlement and 41 percent from bank larceny,[6] and Grogger finds that arrests can explain about two-thirds of the black/white youth employment differential in his sample.[7] In two other studies, however, Grogger finds only a short-lived effect of arrests on youth earnings.[8] Some studies have even found increases in wages after conviction, and explain this as the consequence of

the types of jobs available to those stigmatized, which may pay more initially but offer less chance for advancement.[9]

These articles sought to measure the amount of stigma, rather than to explain its presence. One explanation for stigma is as a taste: people feel moral repugnance for criminals, and choose to incur personal costs rather than interact with them. Under this explanation, an employer would sacrifice profit by refusing to hire someone with a criminal record even though he would have to pay higher wages to a less productive employee with a clean record. This has some plausibility, but appeal to tastes is not necessary to explain stigma. The present article will construct two models, explaining it not as dislike of crime for its own sake but as a rational consequence of the association of criminality with other, directly undesirable, characteristics.

Section II constructs two formal models of the interaction between potential criminals' decisions to commit crimes and employers' decisions to stigmatize detected criminals. Both models show how multiple equilibria could exist, some with low stigma and high crime is low, and others with low stigma and high crime. Section III explores the implications of stigma for public policy in the areas of enforcement, punishment, and disclosure. Section IV concludes.

II. MODELLING STIGMA

The idea to be modelled is that public declaration of a person's criminality makes other people reluctant to interact with him. In the models, this reluctance will take the form of employers paying lower wages to those convicted of crimes.[10]

Two models will be developed, both based on the idea that conviction conveys information about criminality and that employers prefer not to hire criminals, but have no direct taste for stigmatization. In the moral hazard model, all workers will begin with equal marginal products, but anyone who engages in crime becomes less productive. In the adverse selection model, crime will have no effect on productivity, but some workers are less productive regardless of whether they commit crimes, and, for exogenous reasons, these workers have a greater tendency to commit crimes.

It is important that the reader understand this terminology. First, consider "taste for stigmatization." A purely profit-maximizing employer has no taste for stigmatization and will hire workers based solely on the profit he expects to earn from them. If worker X, an active child molester in his personal life, would contribute one dollar per year more to the firm's profit than worker Y, who has less detestable hobbies, the employer would prefer X.

Tastes for stigmatization undoubtedly exist, especially in social relations, but they will not be assumed here. If stigma can be explained without recourse to taste, in fact, one may be able to explain the taste for stigmatization as a rule of thumb based on deeper preferences.

Second, consider "productivity". The mental image this calls to mind is a factory worker making widgets; if he turns out more widgets, his marginal product is higher. Both theoretically and empirically, however, economists use the term to refer to the extra output produced by the firm when a worker is added to its labor force. If he produces ten widgets, but spoils two, steals three, and interferes with the neighboring worker enough to reduce his output by four, our worker's marginal product is not ten, but one. Thus, when the moral hazard model assumes that the productivity of a worker who engages in crime falls, it does not assume that his ability to perform tasks declines.[11] The most obvious link between crime and productivity is employee theft and the precautions needed to avoid it. Dickens, Katz, Lang and Summers cite studies claiming that employee theft costs American business between 15 and 56 billion dollars per year, accounts for between 5 and 30 percent of business failures, and induces spending of 12 billion dollars per year on prevention.[12] Someone with a history of crime has learned criminal techniques, discovered how to fence stolen goods, and overcome the fear and conscience pangs of a first offense. Employers may also reasonably be concerned that he has become more willing to steal time by shirking on the job. In these ways, his productivity falls.

The adverse selection model's assumption that criminality is correlated with low productivity is easier to visualize. One easily quantified link is that criminality is correlated with low intelligence, and intelligence, in turn, is correlated with productivity.[13] The low productivity in the adverse selection model, however, could also derive from the same characteristics as in the moral hazard model— tendencies to steal and to shirk— but tendencies that exist independent of the worker's choice, or which while causing criminality are not caused by it.

Keeping this in mind, let us proceed to the formal models.

A. Moral Hazard and Stigma

The decisionmakers in the model are risk-neutral workers and risk-neutral employers. The workers must decide whether to commit crimes. Criminals are unobserved by employers unless they are caught and convicted, and employers must decide how much to pay convicted and unconvicted workers.

If a worker decides to engage in crime, he is caught and convicted with

exogenous probability $\alpha \in (0,1)$. The direct reward from crime is V and the public penalty from being convicted is P.[14] There is a continuum of workers, so $\theta \in [0,1]$, the proportion that choose crime, is unaffected by any individual's decision. Workers are identical except for a heterogeneity parameter u with cumulative distribution $F(u)$ across the population, where a positive u denotes an individual whose aversion to crime is greater for unmodelled reasons such as moral scruples, lack of skill, or poor criminal opportunities.[15] Let $F'(u) > 0$ for any u, which implies that some people will choose crime no matter how high the penalties and some people will refrain from crime no matter how low the penalties.

Whether a worker has been convicted or not, he offers himself for employment. Crime hurts net productivity. In legitimate employment, the criminal's marginal product is m and the noncriminal's is $m + y > m$. This may be so for a variety of reasons, including employee theft, resistance to authority, and lack of attention to acquiring legitimate skills. Employers compete with each other for workers, but all they observe are convictions, not criminality or marginal product.

In equilibrium, a convicted worker will earn his marginal product of m. A worker whose innocence was known would receive $m + y$, but the category of unconvicted workers pools noncriminals with unconvicted criminals, so the wage for an unconvicted worker, w, will lie in the interval $[m, m + y]$ and depend on the proportion of unconvicted workers believed to be criminals. Fraction α of the θ criminal workers are convicted, leaving proportion $(1 - \alpha\theta)$ of the population unconvicted, which is the denominator for the expected-value expression (1) below. Of the unconvicted $(1 - \alpha\theta)$, amount $(1 - \theta)$ are noncriminal and have marginal product $m + y$, while amount $\theta(1 - \alpha)$ are unconvicted criminals with marginal product m. Hence, the average marginal product in the unconvicted population is

$$
\begin{aligned}
w &= \left(\tfrac{1-\theta}{1-\alpha\theta}\right)(m + y) + \left(\tfrac{\theta(1-\alpha)}{1-\alpha\theta}\right) m \\[2mm]
&= m + \tfrac{1-\theta}{1-\alpha\theta} y.
\end{aligned}
\tag{1}
$$

It can immediately be seen that the wage of the unconvicted worker falls with the amount of crime:

$$
\frac{\partial w}{\partial \theta} = -\left(\frac{1 - \alpha}{(1 - \alpha\theta)^2}\right) y < 0.
\tag{2}
$$

Depending on whether he is criminal or noncriminal, the worker's expected payoff is

$$
\pi_c = (V - \alpha P) + (1 - \alpha)w + \alpha m - u
\tag{3}
$$

or

$$\pi_{nc} = w. \tag{4}$$

The worker will choose to be criminal if $\underline{A}$, the attractiveness of crime, is positive. Using (1), (3), and (4), its value is

$$
\begin{aligned}
A & \equiv \pi_c - \pi_{nc} \tag{5} \\
& = V - \alpha P + (1 - \alpha)w + \alpha m - u - w \\
& = (V - \alpha P) - \alpha \left(\frac{1 - \theta}{1 - \alpha\theta} \right) y - u. \tag{6}
\end{aligned}
$$

Proposition 1 summarizes the interactions of the attractiveness of crime with the other variables in the model.

PROPOSITION 1: The attractiveness of crime is: (a) increasing in the direct reward to crime, V ; (b) decreasing in the personal disutility of crime, u ; (c) decreasing in the criminal penalty, P; (d) decreasing in the productivity damage, y ; (e) decreasing in the probability α of conviction, even if $P = 0$; and (f) increasing in the aggregate crime rate, θ.[16]

Points (a) through (e) are not unexpected. Crime increases with its rewards and falls with its penalties, as in any economic model of crime. Note, however, that although increasing the probability and amount of punishment reduce the attractiveness of crime, the effects of α (the probability of conviction) and $\underline{P}$ (the penalty) diverge. $\underline{P}$ exerts a negative effect only once in equation (6), when it is multiplied by α in the official punishment. α exerts two additional negative effects. If the probability of conviction is high, then (i) the probability of being convicted and stigmatized is higher and (ii) the amount of stigma is greater. Contrary to simpler models of crime, enforcement has more impact than punishment: holding the expected penalty $\underline{\alpha P}$ constant at 2, a value of $\underline{P = 20}$ combined with $\alpha = .1$ would not deter crime as strongly as $\underline{P = 4}$ and $\alpha = 0.5$. Even if $\underline{P = 0}$, the threat of stigma might be sufficient to deter crime by itself.

What is most interesting, however, is part (f): the effect of the general crime rate on the individual who is considering whether to commit crimes. The variable θ representing the proportion of criminals is endogenous; it determines the individual worker's decision, but itself is determined by the decisions of all workers. When θ rises, the wage loss from conviction falls. The payoffs of both criminal and noncriminal workers decline, but the payoff of noncriminal workers declines more. From equations (3) and (4) it can be seen that $\frac{\partial \pi_{nc}}{\partial \theta} = \frac{\partial w}{\partial \theta}$, and $\frac{\partial \pi_c}{\partial \theta} = -(1 - \alpha)(\frac{dw}{d\theta})$. Since $\frac{\partial \pi_c}{\partial \theta} = -(1 - \alpha)\frac{\partial \pi_{nc}}{\partial \theta}$, a high crime rate hurts the noncriminal more than the criminal.

There exist cutoff levels of u such that individuals with heterogeneity

parameters in the interval $[-\infty, \underline{u}]$ will always engage in crime, those in the interval $(\underline{u}, \overline{u})$ will decide based on θ, and those in the interval $[\overline{u}, +\infty]$ always refrain from crime.[17] In this interval, let $\tilde{\theta}(u)$ be the crime level at which an individual of type u is indifferent between crime and noncrime. For $\theta > \tilde{\theta}(u)$ he will choose crime, and for smaller θ he will not. $\tilde{\theta}(u)$ is increasing in $\underline{u}$, and this implies that if $\theta = \tilde{\theta}(u)$, all individuals in the interval $[-\infty, u]$ will choose crime. Figure 1 puts $F(u)$ and $\tilde{\theta}(u)$ on the same diagram. Any intersection (u^*, θ^*) between the curves $\tilde{\theta}(u)$ and $F(u)$ will be an equilibrium. At θ^*, the marginal criminal, with utility parameter $\underline{u}^*$, will be indifferent about choosing crime because $A(u^*, \theta^*) = 0$, while the $F(u^*)$ individuals with lower levels of u will choose crime and the $(1 - F(u^*))$ with higher levels will refrain from crime.
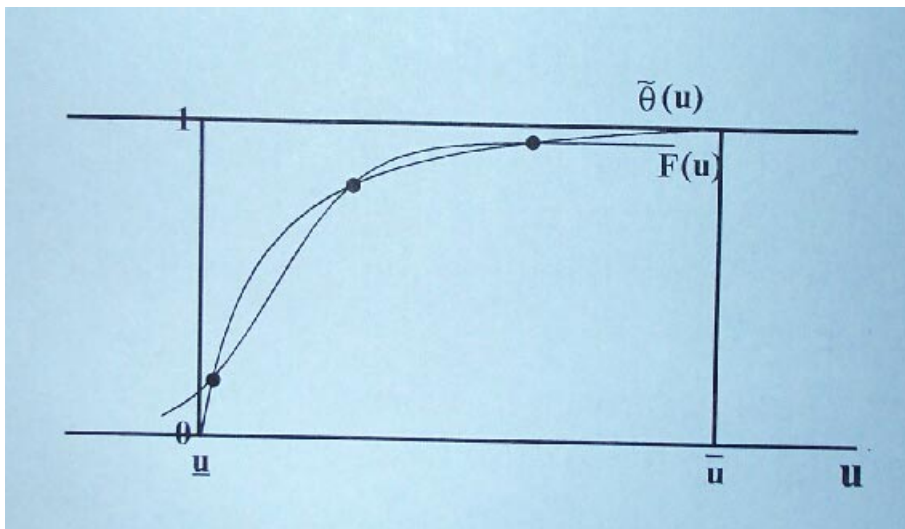
Proposition 2 says that the assumptions of the model permit multiple equilibria of this kind.

PROPOSITION 2: Depending on the distribution of the taste for crime, $F(u)$, there may exist multiple equilibria. If there are three equilibria, with crime levels $\theta^- < \theta^* < \theta^+$, then the two outer equilibria are stable and the middle one is unstable. The equilibria can be pareto-ranked, with lower crime levels being superior.

Proof. (i) Existence. An equilibrium is at an intersection of $\tilde{\theta}(u)$ and $F(u)$. $\tilde{\theta}(\underline{u}) = 0$ and $\tilde{\theta}(\overline{u}) = 1$ by definition of $\underline{u}$ and $\overline{u}$. That $\tilde{\theta}(u)$ is increasing and continuous can be seen as follows. $\tilde{\theta}$ is found by setting A equal to zero and solving for θ in equation (6):

$$V - \alpha P - \frac{1 - \theta}{1 - \alpha\theta}\alpha y - u = 0, \tag{7}$$

which implies that

$$\tilde{\theta} = \frac{V - \alpha P - \alpha y - u}{\alpha(V - \alpha P - y - u)}. \tag{8}$$

The derivative of (10) with respect to $\underline{u}$ exists and is

$$\frac{\partial \tilde{\theta}}{\partial u} = \frac{-1}{\alpha(V - \alpha P - y - u)} + \frac{\alpha(V - \alpha P - \alpha y - u)}{\alpha^2((V - \alpha P) - y - u)^2} = \frac{y(1 - \alpha)}{\alpha(V - \alpha P - y - u)^2}, \tag{9}$$

which is positive because $\alpha \in (0, 1)$.

Because $\tilde{\theta}(u)$ is continuous, increasing, and takes every value between 0 and 1, and F is nondecreasing, continuous, and restricted to values between 0 and 1, there must be some u^* at which $F(u^*) = \tilde{\theta}(u^*)$. Thus, an equilibrium exists. The curve $\underline{F}$ might intersect $\tilde{\theta}(u)$ at more than one point, generating the multiple equilibria of Figure 1, or it might intersect just once.

(ii) Stability. An equilibrium θ is stable with respect to a dynamic process if for arbitrarily small ϵ and an initial state $(\theta + \epsilon)$ or $(\theta - \epsilon)$, the limit of the dynamic process is θ. The simplest dynamic process is myopic: "In period $\underline{t}$, individuals make their decisions as if they believe that θ_t will equal θ_{t-1}."

An equilibrium's stability depends on whether F cuts $\tilde{\theta}(u)$ from above or below. If F cuts $\tilde{\theta}(u)$ from above (that is, if $F(u^* - \epsilon) > \tilde{\theta}(u^* - \epsilon)$ and $F(u^* + \epsilon) < \tilde{\theta}(u^* + \epsilon)$), then the equilibrium is stable. Suppose that these inequalities are true and the system starts at $\theta' < \theta^*$ where $u = u^* - \epsilon$. The amount of crime will increase, because $F(u) = \theta'$, which is greater than $\tilde{\theta}(u)$,the amount of crime that induces individual u to undertake crimes. If F intersects $\tilde{\theta}(u)$ from above, then F must have a gentler slope than $\tilde{\theta}(u)$ at $u^* - \epsilon$, so the increase in crime will not overshoot u^*, and myopic dynamics will converge at u^*.

If F does not cut $\tilde{\theta}(u)$, but rather intersects it at the extreme values of $\overline{u}$ or $\underline{u}$, then the equilibrium is still stable. The previous paragraph's argument still applies to dynamics starting from values of u nearer 0 than u^*, and if the system starts at a more extreme value of u, even myopic dynamics instantly lead back to u^*. If there is a single equilibrium, then $F(u)$ either intersects $\tilde{\theta}(u)$ at an extreme value, in which case the same argument shows it is stable, or $F(\underline{u}) > 0$ and $F(\overline{u}) < 1$. But if this is the case and F is continuous, then F, starting greater than $\tilde{\theta}(u)$ and ending smaller than $\tilde{\theta}(u)$, must cut $\tilde{\theta}(u)$ from above, and the equilibrium is stable. If there are three equilibria, then any equilibrium at $\underline{u}$ or $\overline{u}$ is an outer equilibrium, and is stable by the same argument. That argument also shows that the smallest equilibrium must either be at $\underline{u}$ or (given that $\tilde{\theta}(u)$ is upward sloping) at a point where $F(u)$ cuts $\tilde{\theta}(u)$ from above. But if F cuts $\tilde{\theta}(u)$ from above at the first equilibrium,

then it must cut from below at the middle equilibrium, u^*. And if it cuts from below at u^*, then for slightly larger u, $F(u)$ lies above $\tilde{\theta}(u)$, and it must cut $\tilde{\theta}(u)$ from above at the final equilibrium. Hence the two outer equilibria are stable, and the middle equilibrium is not.

(iii) Optimality. Even from the point of view of the potential criminals, the high-crime equilibrium is dominated by the low-crime equilibrium. Inequality (2) shows that w falls in θ, and equations (3) and (4) show that w is a component of the payoffs of both the criminal and the noncriminal, so the high-crime equilibrium has lower payoffs for all. Q.E.D.

Proposition 2 establishes the possibility of multiple, pareto-ranked, stable equilibria. Every individual, whether his particular tastes lead him to be criminal or noncriminal, prefers the low-crime equilibrium, in which stigma has a strong effect and convicted criminals receive large cuts in their wages. This is less paradoxical when it is rephrased: every individual prefers the equilibrium in which lack of a criminal record is rewarded by a wage premium. The stigma punishment and the wage-premium reward are equivalent. What matters is that there be a wedge between the wages of the convicted and the unconvicted.

How plausible are multiple equilibria? If an increase in crime is to be largely explained by a reduction in stigma, it must also be true that a significant proportion of the population– or at least of subpopulations such as young males– has become criminal. The shift in the proportion of criminals need not be from 0 percent to 100 percent, but if it is merely from 1 percent to 5 percent the effect on average productivity, and thus on stigma, will be small.

Criminality is indeed very common among young males.[18] Ball, Ross and Simpson found that as early as 1960, 20.7 percent of the boys and 5.3 percent of the girls in Lexington, Kentucky had appeared in juvenile court.[19] Tillman examined a comprehensive set of arrest records to discover the probability of being arrested for Californians who were 18 in 1974, and found that 34 percent of the white males and 66 percent of the black males were arrested (41 percent of the black males for a felony).[20] Most arrests are for public order offenses such as drunk driving and disorderly conduct and charges are not pressed, but the number of men convicted of crimes is also remarkably high. In 1993, the number of men in jail or prison equalled 1.9 percent of the male labor force, and the number on probation or parole added a further 4.7 percent.[21] Thus, the number of men currently being punished in one way or another was 6.6 percent of the labor force. These figures, moreover, are for the entire male population. Of men aged eighteen to thirty-four, the fraction under supervision of the courts was 11 percent, and for black men in that

age group it was 37 percent.[22] Since not all criminals are caught and not all those caught are in prison simultaneously, the total number of past and present criminals must be astonishingly high. Clearly, multiple equilibria cannot be ruled out on the grounds that too few people engage in crime to seriously affect the quality of the labor force.

B. Adverse Selection and Stigma

The moral hazard model captures the channels by which stigma operates when crime reduces productivity. Stigma can be effective, however, even when crime does not reduce productivity. In that case, whether an employee committed crimes in the past would not affect his wage in a world of perfect information, but under imperfect information, employers might use criminality as a proxy for low productivity.

To model this, let m be the marginal product of the low-ability workers, who always commit crimes and who form proportion $\overline{\theta}$ of the population. Let $m + y$ be the marginal product of high-ability workers, who choose whether or not to commit crimes and who form proportion $1 - \overline{\theta}$ of the population. The total proportion of criminal workers is $\theta > \overline{\theta}$. A criminal is caught and convicted with probability α, and crime has no effect on productivity. Let us assume, for simplicity, that there is no other worker heterogeneity of the kind that $F(u)$ represented in the moral hazard model.

In equilibrium, the wage for a convicted worker is not necessarily m, as in the moral hazard model, because high- ability workers might be convicted too. Low and high-ability workers are convicted at the same rate, so all that matters is the relative proportion in the criminal population. The wage for the convicted is

$$w_c = \left(\frac{\overline{\theta}}{\theta}\right) m + \left(\frac{\theta - \overline{\theta}}{\theta}\right) (m + y), \tag{10}$$

which equals m only if $\theta = \overline{\theta}$.

The unconvicted population is composed of unconvicted criminals with low ability (proportion $\overline{\theta}(1 - \alpha)$), unconvicted criminals with high ability $((\theta - \overline{\theta})(1 - \alpha))$, and noncriminals with high ability $(1 - \theta)$, a total probability mass of $1 - \alpha\theta$. The unconvicted wage is therefore

$$w = \left(\frac{\overline{\theta}(1 - \alpha)}{1 - \alpha\theta}\right) m + \left(\frac{1 - \overline{\theta} - \alpha(\theta - \overline{\theta})}{1 - \alpha\theta}\right) (m + y). \tag{11}$$

The adverse selection model has two equilibria: a pooling, high-crime equilibrium in which high-ability workers choose crime and the unconvicted

wage equals the convicted wage; and a separating, low-crime equilibrium in which high-ability workers refrain from crime and conviction carries stigma. In the low-crime equilibrium, only low-ability workers commit crimes, and convicts are paid the low-ability wage. High-ability people refrain from crime, because they do not want to risk being pooled with the low- ability convicts. The unconvicted are paid a wage between the low- and high- ability wages, because low-ability criminals who are not caught are indistinguishable from high- ability workers. In the high-crime equilibrium, everyone commits crimes, and the wage for the convicted and the unconvicted is the same.

Formally, in the low-crime equilibrium, none of the high-ability workers choose crime. This means that $\theta = \overline{\theta}$,

$$w_c = m, \tag{12}$$

and

$$w = \left(\frac{\overline{\theta}(1 - \alpha)}{1 - \alpha\overline{\theta}}\right) m + \left(\frac{1 - \overline{\theta}}{1 - \alpha\overline{\theta}}\right)(m + y). \tag{13}$$

In the high-crime equilibrium, all of the high-ability workers choose crime. This means that $\theta = 1$,

$$w_c = \overline{\theta}m + (1 - \overline{\theta})(m + y), \tag{14}$$

and

$$w = \left(\frac{\overline{\theta}(1 - \alpha)}{1 - \alpha}\right) m + \left(\frac{1 - \overline{\theta} - \alpha(1 - \overline{\theta})}{1 - \alpha}\right)(m + y)$$

$$= \overline{\theta}m + (1 - \overline{\theta})(m + y). \tag{15}$$

The wage is the same for both convicted and unconvicted workers in the high-crime equilibrium. Neither equilibrium is Pareto- dominant, in contrast to the moral hazard model. Low-ability workers prefer the high-crime equilibrium, but high-ability workers prefer the low-crime equilibrium.

The moral hazard and adverse selection models make similar predictions, except that a move from low to high crime leaves the wage for the convicted unchanged in the moral hazard model and raises it in the adverse selection model. This is because in the adverse selection model, the average wage is independent of the number of criminals. As criminality increases, the wage of the unconvicted falls, but the wage of the convicted rises. The biggest difference is perhaps in the welfare implications, since in the adverse selection model the cost of high crime is limited to the crime itself rather than to ill effects on worker productivity.

Both the adverse selection model and the moral hazard model are based on asymmetric information– the worker knows his productivity level, but

the employer does not. In this, they are different from another phenomenon that one might associate with the term "stigma": reputation effects of the kind usually modelled as repeated games of symmetric information. If a prisoner's dilemma is repeated an infinite number of times with sufficiently little discounting, the two players may each choose to cooperate for fear that a betrayal would lead to a cessation of cooperation by the other player. The reputation model of Klein and Leffler relies on essentially the same reasoning: a firm produces high-quality products because if it ever betrays consumers with a low-quality product, they will cease buying.[23] David Hirshleifer and myself have shown how a model of this kind can support costly ostracism: members of a group can be given the incentive to expel an offending member even if his presence would add to the group's wealth.[24] This distrust of a player who has deviated from cooperation seems especially apt for situations where a criminal has offended against the person who stigmatizes him—stealing from his employer, for example, who thereupon fires him. Reputation models are based on self-fulfilling expectations, but they have different implications from the stigma models of this article. In a reputation model, it is not a person's background type or his past criminality that makes him an undesirable employee; it is his belief that the employer does not trust him, which in turn may be based on the employer's knowledge of the employee's criminal background. This distrust of the employer is self-fulfilling but arbitrary, and a reputation model could equally well assume that employers distrust _non-_ criminals, who in turn become bad employees because of that distrust. In the stigma models, in contrast, the employee's productivity is independent of the attitude of his current employer, which affects only how much he is paid.

III. STIGMA AND PUBLIC POLICY

A. The Government's Choice of the Probability of Conviction

In the standard economic model of crime, only the expected penalty matters for deterrence, and the division between punishment and probability of conviction is important only to the government's expense of punishment. In the stigma model, the probability of conviction, α, has a double deterrent effect, operating via not only the public punishment $\underline{P}$, but private stigma. Even if $\underline{P} = 0$, if stigma is sufficiently great, crime is deterred.[25]

Paradoxically, the productivity loss from crime can be beneficial to the potential criminal and to society, because it permits a low- crime equilibrium to exist even when official penalties are low. The productivity loss helps to explain the lower crime rates of the affluent, since for many well-paying jobs a large productivity loss is plausible. Lott has shown empirically that a larger
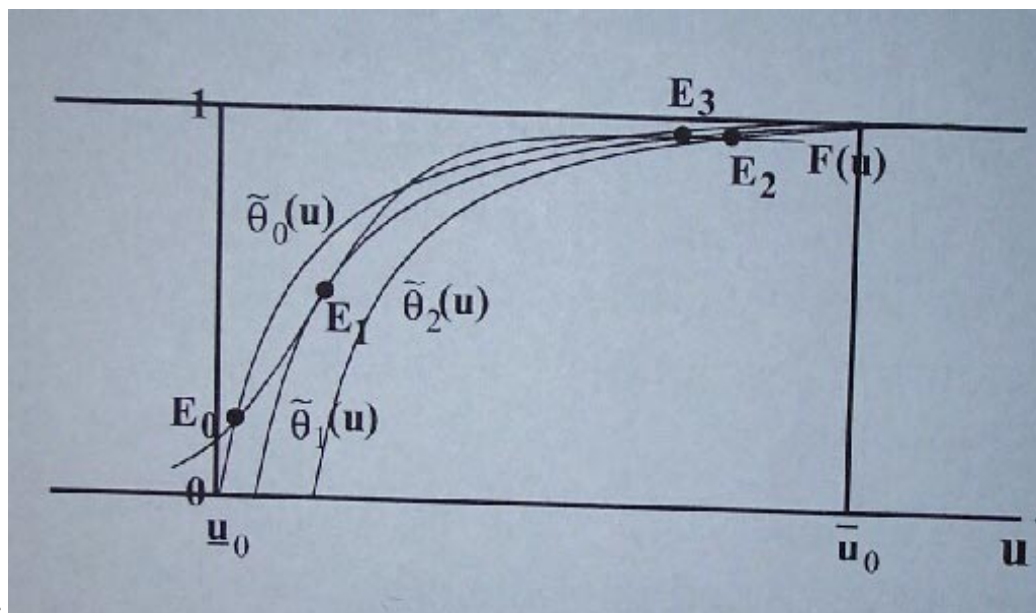
portion of the punishment for a wealthier person is indeed in the form of wage loss.[26] Facing a heavier penalty, he is more strongly deterred. Stigma may also help explain why crime rates are so high among the young. For reasons unrelated to crime, young people are less likely to be employed, and therefore less likely to suffer immediate economic stigma if caught. Although the participation rate for males aged 16-19 is only 53 percent , it rises to 94 percent for males aged 25 to 34.[27] The situation is self-reinforcing, since employers are more reluctant to hire the young if they are disproportionately criminal.

The probability of conviction is thus a more powerful policy tool than the criminal penalty, when stigma is effective. What conviction probability is optimal? That depends on the relative costs of enforcement and crime, on which equilibrium is in effect, and on the likelihood of random shocks to individual tastes for crime as I will now explain.

Figure 1 showed $F(u)$, the distribution of individual aversions to crime, and $\tilde{\theta}(u)$, the critical levels of crime that induce different individuals to engage in crime. In Figure 2, a reduction in enforcement (the probability α or the size $\underline{P}$ of punishment) shifts down $\tilde{\theta}(u)$ from $\tilde{\theta}_0(u)$ to $\tilde{\theta}_1(u)$ If the system begins at E_0, then whether enforcement should be increased or reduced depends on its cost compared to the costs of crime– productivity loss, victim precautions, and so forth. If crime is costly compared to enforcement, then $\tilde{\theta}(u)$ should be shifted up by increasing the amount of enforcement. If enforcement is more costly, then $\tilde{\theta}(u)$ should be shifted down. Figure 2 shows the effect of reduced enforcement: $\tilde{\theta}(u)$ shifts to $\tilde{\theta}_1(u)$ and the equilibrium moves smoothly from E_0 to higher crime at E_1.

FIGURE 2 GOES HERE

For a wide range of costs of enforcement, E_1 will be the equilibrium re-

sulting from the optimal c
of enforcement level. If enforcement is reduced any further, $\tilde{\theta}(u)$ shifts to
$\tilde{\theta}_2(u)$ and the equilibrium shifts discontinuously to much higher crime at E_2.
E_1 is the equilibrium with the lowest level of enforcement that still enables
private stigma to effectively supplement public punishment.

Equilibrium E_1, however, is not robust to small shocks in α, $\underline{P}$, and $F(u)$.
If enforcement dips slightly, or individuals become less averse to crime, then
the low-crime equilibrium disappears, and crime increases discontinuously.
A small shock can be drastically multiplied. Since the expectations that
maintain the low- crime equilibrium are a form of valuable social capital, the
presence of random shocks would make a higher level of enforcement optimal
than would otherwise be the case, and the optimal expected amount of crime
would be less than $\tilde{\theta}(u)$.

The optimal level of enforcement also depends on which equilibrium is in
effect. Suppose that enforcement has fallen enough that E_2 is the equilibrium.
If enforcement increases, the $\tilde{\theta}(u)$ curve returns to $\tilde{\theta}_1(u)$, but the equilibrium
does not return to drastically lower crime at E_1, but to slightly lower crime
at E_3. Thus, although enforcement levels resulting in $\tilde{\theta}_1(u)$ may be optimal
starting from $\tilde{\theta}_0(u)$ or $\tilde{\theta}_1(u)$, if the system begins at $\tilde{\theta}_2(u)$ a lower level of
enforcement may be optimal. If crime is low, long jail sentences may be
optimal to maintain stigma, but if crime is high, and stigma has ceased to
work, the authorities should give up and become more lenient. The optimal
enforcement effort can actually fall as crime increases. The stigma model
also suggests that a "big push" would be the most effective way to reduce
crime: it may be worth investing resources to push the system back to the
low-crime equilibrium, even if it is not worthwhile trying to ameliorate the

high-crime equilibrium.

The other side of the coin is that an increase in crime induced by misguided policies may be very difficult to reverse. This may help explain the puzzle mentioned in the Introduction: that criminal penalties seem less effective for deterrence in the 1980's than in the 1960's. The stigma model suggests the following story. In 1960, the United States was at a low-crime equilibrium, in which a combination of public punishment and private stigma deterred crime. A number of things then happened to make crime more attractive, including perhaps a general decline in the penalties of conscience (a shift rightwards of $F(u)$ in Figure 2) and certainly a lenient government policy (a decline in $\underline{P}$ and α, which would shift the $\tilde{\theta}(u)$ curve downwards)[28].Eventually there existed just one equilibrium, with high crime, and the value of θ moved towards it as expectations changed.

In the 1970's and 1980's, horrified public opinion forced a rise in the imprisonment rate. According to the stigma model, increasing the punishment rate would shift $\tilde{\theta}(u)$ up again, reducing the amount of crime slightly, but if the crime rate in 1970 were greater than the middle, unstable, equilibrium, the adjustment process would continue to push the crime rate up to the high-crime equilibrium. Thus, the crackdown reduced crime per youth slightly, but crime soon rose again, albeit more slowly. The increase in punishment could not overcome the loss of stigma. Although arrest rates did not increase much during the 1980s (from 515 per 100,000 in 1980 to 558 in 1992[29], imprisonment rates rose sharply, which caused a decline in crime in the early 1980s. The decline was small compared to the increase in the 1960s, because the tough policy of the 1980s was not tough enough to restore stigma.

Support for the stigma explanation is provided by changes in the pattern of arrest rates by age category. Young people became more criminal, and older people less criminal.[30] This is especially curious because someone who was 21 in 1971 had a probability of arrest much greater than that of his uncle who was 21 in 1961, but by the time he reached age 35 his probability of arrest was lower than his uncle's at the same age.

An explanation is that stigma can decline for a subpopulation such as young men even if it retains its strength for the middle-aged. The young have not yet established a reputation for productivity in the labor market and their employers are at more of an informational disadvantage. As a result, the decline in stigma had a disproportionate effect on youth crime.[31] The increase in official punishment since 1971, on the other hand, has affected both young and old, so arrests of older people increased less, or even declined. Grogger notes that between 1973 and 1988, real wages paid to young men who worked full-time fell 23 percent , which would more than explain the

increase in youth arrest rates over that period according to the elasticity of crime with respect to wages that he estimates.[32] The stigma model suggests that causality went both ways, and real wages fell because crime increased.

B. The Advantages of Stigma as a Punishment

One of the oldest issues in the economics of crime is how society can deter crime efficiently. Imprisonment is costly, and Becker has suggested that fines be used wherever possible because they are transfers rather than social costs.[33] If the fine is large and the probability of detection small, the expected penalty can be large enough to yield deterrence at a low social cost. This policy has well-known practical problems, of which the most important is the inability of criminals to pay substantial fines. High fines also raise the concern that the government may be tempted to prosecute the innocent for the sake of revenue.

Stigma avoids these problems. Although many people have little liquid wealth, the market value of most people's future labor rents is substantial. Stigma is like a fine drawn on those future rents, a fine which can be collected regardless of the criminal's present wealth. Since it is the private sector that imposes the punishment, stigma is neither costly to the government, like imprisonment, nor revenue-raising, like fines, so neither concern distorts the government's decision.

The main disadvantage of stigma is perhaps that its effectiveness diminishes for recidivists. Stigma is a cheap and efficient punishment, but only for someone with a reputation to lose. The stigma from a first conviction is greater than from subsequent convictions, and after enough convictions the marginal effect is negligible. To achieve a given level of deterrence or retribution, the fine or jail term for the first offense should be much smaller than for the subsequent offenses.[34]

Stigma shares with fines the advantage of deterring the criminal without creating real costs, because it transfers wealth from the criminal to the rest of society. Stigma actually increases efficiency, because allocative efficiency increases as information is disclosed. The stigma from automobile speeding, for example, is that the offender will pay more for automobile insurance after being identified as a fast driver with a disdain for regulations. This comes closer to matching the social cost of the offender's driving with the private cost to himself. The effect in the labor market is similar. Prior to his conviction, the criminal's labor is overvalued in the market. His loss of income after stigmatization is a gain for noncriminal workers who would otherwise be pooled with him and paid less than their marginal products so he could be paid more.[35]

This benefit of stigma is different from the conventional functions of punishment— deterrence, incapacitation, rehabilitation, and retribution. Stigma has advantages as a deterrent, and may even serve to incapacitate the criminal by removing him from jobs that would give him opportunities for crime, but in disclosing information stigmatization serves a distinctly different function. Even if stigma had no effect on the amount of crime, it would improve efficiency.

C. Publicizing Government Records

Because stigmatization is distinct from deterrence, courts need to convey accurate information to the public, rather than just inflicting the appropriate penalty. For deterrence, it may not matter if the court declares someone guilty of counterfeiting rather than his actual crime of burglary, so long as the penalty is appropriate for burglary. For stigmatization, however, the exact charge is important, because different kinds of people commit different crimes.

This points to a danger in using plea bargaining to reduce the cost of prosecution. In a common type of plea bargain, the accused pleads guilty to a crime milder than that for which the prosecutor has good but not overwhelming evidence. From the point of view of stigmatization, it would be much better for the plea bargain to take the form of a guilty plea to the original crime, but with a recommendation of a reduced sentence. The public penalty would be the same, but stigma could be more accurately applied.

The social utility of stigma is also relevant to the question of whether criminal records should be open to the public. Court dockets are open as a matter of constitutional right, and daily police arrest blotters are traditionally open, but the availability of records filed by name varies state by state.[36] State legislatures have passed a wide variety of statutes ranging from Florida's completely open records to Illinois' restriction of access to providers of child care, volunteer organizations associated with children, detective agencies, security-guard organizations, schools, and liquor-license holders. Moreover, juvenile records are often kept secret even when adult records are not.

In contrast to the general trend in American law of valuing the individual's privacy over other people's accurate knowledge about him, there has been a surge of legislation in the 1990's designed to stigmatize sex offenders. As of 1995, thirty-eight states had sex offender registration laws, which in their usual form require anyone convicted of rape or child molesting to register with the police chief in the town in which they live.[37] Older laws limited

the public's access to the police registry, but it is now common to allow not only access, but convenient access. In California, a "900" telephone number exists for inquiries about particular individuals by name, street address, or other indexing information, and Louisiana law takes publicity even further, for parolees at least, by authorizing the parole board to require the use of bumper stickers or labelled clothing.[38] The motivation is clearly not to increase the magnitude of the punishment, but to allow other people to make use of the information in their dealings with the ex-convict.

The argument for keeping criminal records secret is that by preventing discrimination against workers with criminal pasts it gives them higher wages in legitimate employment and greater motivation for a fresh start. This is sometimes joined to the argument that employers are unreasonably prejudiced against workers with criminal records, because criminality is not associated with productivity. This argument is weak both because it offends common sense (would you really be indifferent about whether your warehouseman had a background as a burglar?) and because it presumes that the persons asserting it know more about ex-criminals' productivity in particular jobs than employers do.[39] Even if the argument were valid, however, and stigma were based on mistaken beliefs about productivity, it would not be conclusive, because stigma would still be useful as a punishment. Stigma based on mistaken beliefs would be a costly punishment because of its distorting effect on the labor market, more like imprisonment than fines, but it might still be optimal.

A stronger argument against stigma is based on possible positive externalities from employing criminals. If employers were forbidden access to criminal records, they would overestimate the convicted criminal's productivity and pay him a higher wage. The direct effect would be to hurt allocative efficiency, since employers would pay a uniform wage which would exceed the criminal workers' marginal product and be less than the noncriminal workers' marginal products. At the higher wage, however, more criminals would choose to be employed in legitimate jobs, and this would raise the opportunity cost of crime.[40] This social benefit does not figure in the employer's calculations, so it may be socially beneficial to keep criminal records secret.[41] The tradeoff is between the beneficial effect of secrecy on recidivism and the harmful effects on deterrence of first crimes and on allocative efficiency.

Against this benefit must be set the disadvantage that lack of stigma increases the incentive for crime in the first place. No policy that tries to induce the convicted criminal to refrain from crime by increasing the benefits of legitimate work can escape this incentive problem, but not all policies create the allocative distortions of secret records. Those distortions could be avoided by tackling the externality problem directly, keeping records open,

but subsidizing the wage of ex-criminals. The ex-criminal would then become employed, but would be better matched with jobs; the former embezzler could be hired as a schoolteacher, the former child molester as a bookkeeper. A wage or training subsidy would weaken the deterrence effect of stigma, but it would not distort the labor market.

IV. CONCLUDING REMARKS

Since Becker's seminal article in 1968, economists studying crime have focussed on how the probability and severity of punishment deters a potential criminal bent on maximizing his utility. This approach emphasizes the criminal justice system, not the moral disapproval of the society in which the system operates. Reversing the usual pattern, economists stress the role of the government, and sociologists stress the private sector.

The private sector, however, unofficially punishes known criminals by stigmatizing them. Once the criminal's behavior becomes known, other individuals become more reluctant to interact with him. This private reluctance may be as powerful a disincentive to crime as public punishment. The government remains important, but only as a source of detection and provision of reliable information about individual criminality. Government stigmatization is extremely important, but its purpose is really to provide the private sector with the raw materials for the true punishment.

The models used in this paper described economic stigma, a reduction in the wage employers are willing to pay someone with a criminal record either because engaging in crime reduced productivity (the moral hazard model) or correlated with low productivity for other reasons (the adverse selection model). Social stigma could be modelled similarly, as a reduction in the concessions that potential friends or spouses are willing to make to a convicted individual for the privilege of social interaction with him. Whatever its nature, the stigma of a criminal record depends on the informativeness of that record, and thus on the likelihood that someone without a conviction is nonetheless criminal. It was shown that this generates multiple equilibria, because if crime is sufficiently prevalent, a criminal record loses its informativeness and thus its stigmatizing effect.

BIBLIOGRAPHY

Andvig, Jens & Karl Moene. "How Corruption May Corrupt." Journal of Economic Behavior and Organization 13 (1990): 63-76.

Ball, John, Ross, Alan, and Simpson, Alice. "Incidence and Estimated Prevalance of Recorded Delinquency in a Metropolitan Area." American Sociological Review 29 (1964): 90- 93.

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

Bedarf, Abril. "Comment: Examining Sex Offender Community Notification Laws." Calif. L. Rev. 83 (1995): 885.

Bureau of the Census, U.S. Dept. of Commerce. Historical Statistics of the United States: Colonial Times to 1970. White Plains, New York: Kraus International Publications, 1989 (reprint).

Bureau of Justice Statistics, U.S. Dept of Justice, Technical Appendix, Report to the Nation on Crime and Justice, Second Edition, NCJ-112011, July 1988.

Bureau of Justice Statistics, U.S. Dept of Justice. "Public Access to Criminal History Record Information." NCJ-111458, November 1988.

Bureau of Justice Statistics, U.S. Dept of Justice. "Use and Management of Criminal History Record Information: A Comprehensive Report," NCJ-143501, November, 1993.

Bureau of Labor Statistics, U.S. Dept of Labor. Sourcebook of Criminal Justice Statistics, 1988.

Bureau of Labor Statistics, U.S. Dept of Labor. Handbook of Labor Statistics, 1989.

Bushway, Shawn, Daniel Nagin & Lowell Taylor. "The Stigmatic Impact of Criminal Records on Legitimate Employment." Working paper, Heinz School, Carnegie Mellon University, May 1995.

Paul Cassell. "Miranda's Social Costs: An Empirical Reassessment." Nw. U. Law. Rev. 90 (1996): 387.

Dickens, William, Katz, Lawrence, Lang, Kevin, and Summers, Lawrence. "Employee Crime and the Monitoring Puzzle." Journal of Labor Economics 7 (1989): 331-347.

Ehrlich, Isaac. "Participation in Illegitimate Activities: A Theoretical

and Empirical Investigation." Journal of Political Economy 81 (1973): 521-65.

Freeman, Richard. "Crime and the Employment of Disadvantaged Youth," in Adele Harrell and George Peterson, eds., Drugs, Crime and Social Isolation: Barriers to Urban Opportunity, Washington: Urban Institute Press, 1992

Freeman, Richard. "The Labor Market." James Q. Wilson and Joan Petersilia, Eds. Crime, San Francisco, ICS Press, 1995, pp. 171-191.

Freeman, Scott, Grogger, Jeffrey and Jon Sonstelie. "The Spatial Concentration of Crime." Working paper, Dept of Economics, University of California, Santa Barbara, July 1989.

Glaeser, Edward, Bruce Sacerdote and Jose Scheinkman, Crime and Social Interactions, Quarterly Journal of Economics 111 (May 1996): 507-548 .

Grogger, Jeffrey. "Arrests, Persistent Youth Joblessness, and Black-White Employment Differentials." Review of Economics and Statistics 74 (1992): 100-106.

Grogger, Jeffrey. "Market Wages and Youth Crime," working paper, Dept. of Economics, University of California, Santa Barbara, California, February 1994.

Grogger, Jeffrey. "The Effect of Arrest on the Employment and Earnings of Young Men." Quarterly Journal of Economics 90 (1995): 51-72.

Herrnstein, Richard. "Criminogenic Traits." James Q. Wilson and Joan Petersilia, Eds. Crime, San Francisco, ICS Press, 1995, pp. 39-64.

Hirshleifer, David & Eric Rasmusen "Cooperation in a Repeated Prisoner's Dilemma with Ostracism." Journal of Economic Behavior and Organization 12 (1989): 87-106.

Karpoff, Jonathan & John Lott. "The Reputational Penalty Firms Bear from Committing Criminal Fraud." Journal of Law and Economics 36 (1993): 757-802.

Klein, Benjamin & Keith Leffler. "The Role of Market Forces in Assuring Contractual Performance" Journal of Political Economy 89 (1981): 615-41.

Lott, John. "The Effect of Conviction on the Legitimate Income of Criminals." Economics Letters 34 (1990): 381-385.

Lott, John. "An Attempt at Measuring the Total Monetary Penalty from Drug Convictions: The Importance of an Individual's Reputation." Journal of Legal Studies 21 (1992): 159-188.

Lott, John. "Do We Punish High-Income Criminals Too Heavily?" Economic Inquiry 30 (1992): 583-608.

Lui, Francis. "A Dynamic Model of Corruption Deterrence." Journal of Public Economics 32 (1986): 215-236.

Miller, Neal. State Laws on Prosecutors' and Judges' Use of Juvenile Records." Bureau of Justice Statistics, U.S. Dept of Justice, NCJ 155506, 1995.

Murray, Charles. Losing Ground: American Social Policy 1950-1980. New York: Basic Books, 1984.

Nagin, Daniel & Joel Waldfogel. "The Effects of Conviction on on Income Through the Life Cycle." NBER Working Paper No. 4551, November 1993.

Nagin, Daniel & Joel Waldfogel. "The Effects of Criminality and Conviction on the Labor Market Status of Young British Offenders." International Review of Law and Economics, 15 (1995): 109-126.

Posner, Richard. "Optimal Sentences for White–Collar Criminals." American Criminal Law Review 17 (1980): 409-418.

Rasmusen, Eric. "An Income-Satiation Model of Efficiency Wages." Economic Inquiry 30 (1992): 467-478.

Sah, Raj. "Social Osmosis and Patterns of Crime." Journal of Political Economy 99 (1991): 1272-1295.

Schrag, Joel & Suzanne Scotchmer. "Crime and Prejudice: The Use of Character in Evidence in Criminal Trials." Journal of Law, Economics, and Organization 10 (1994): 319-342.

Schrag, Joel & Suzanne Scotchmer. "The Self-Reinforcing Nature of Crime." Working paper, Graduate School of Public Policy, University of California, Berkeley, May 1994.

Tillman, Robert. "The Size of the 'Criminal Population': The Prevalence and Incidence of Adult Arrest." Criminology 25 (1987): 561-579.

U.S. Dept. of Commerce, Bureau of the Census. Statistical Abstract of the United States. Washington: Superintendant of Documents, U.S. Government Printing Of-

fice. Annual.

Visher, Christy and Roth, Jeffrey. "Participation in Criminal Careers." In Blumstein, Alfred, Cohen, Jacqueline, Roth, Jeffrey, and Visher, Christy, editors, <u>Criminal Careers and "Career Criminals", Volume 1</u>. Washington: National Academy Press, 1986.

Waldfogel, Joel. "Does Conviction Have a Persistent Effect on Income and Employment?" <u>International Review of Law and Economics</u> 14 (1994): 103-119.

FOOTNOTES

1. The seminal article is: Gary Becker, Crime and Punishment: An Economic Approach, 76 Journal of Political Economy 169 (1968).

2. Time series data on crime is surprisingly scattered. "Reported crime" is FBI Index crime here, from the crime rate in Charles Murray, Losing Ground: American Social Policy 1950-1980 (1984) , Table 18, and the population in U.S. Dept. of Commerce, Bureau of the Census, Statistical Abstract of the United States, Washington: Superintendant of Documents, U.S. Government Printing Office, 1989 at 2; and 1990, at 300. "Prisoners" refers to people with sentences of at least one year in state and federal courts, from Bureau of the Census, U.S. Dept. of Commerce, Historical Statistics of the United States: Colonial Times to 1970, White Plains, New York: Kraus International Publications, 1989 (reprint), table H1138; the 1984 and 1993 Statistical Abstracts, at 325 and 343. The 1990 figure uses the number of state prisoners multiplied by 1.062, one plus the 1989 ratio of federal to state prisoners. The number of youths aged 16 to 24, an obvious explanation for the crime increase, only rose by 90 percent during this period. Bureau of Labor Statistics, U.S. Dept of Labor, Handbook of Labor Statistics, 1989, at 13. For additional evidence on the increased propensity to crime, see Richard Freeman, The Labor Market, in James Q. Wilson and Joan Petersilia, Eds. Crime, 1995, pp. 171-191.

3. The idea of the overload theory is mentioned as early as Isaac Ehrlich, Participation in Illegitimate Activities: A Theoretical and Empirical Investigation, 81 Journal of Political Economy 521 (1973), and can be found formalized in Francis Lui, A Dynamic Model of Corruption Deterrence, 32 Journal of Public Economics 215 (1986); Jens Andvig & Karl Moene, How Corruption May Corrupt, 13 Journal of Economic Behavior and Organization 63 (1990); Scott Freeman, Jeffrey Grogger & Jon Sonstelie, The Spatial Concentration of Crime, Working paper, Dept of Economics, University of California, Santa Barbara, July 1989; Raj Sah, Social Osmosis and Patterns of Crime, 99 Journal of Political Economy 1272 (1991); and Joel Schrag & Suzanne Scotchmer, The Self-Reinforcing Nature of Crime, Working paper, Graduate School of Public Policy, University of California, Berkeley, May 1994.

4. Recent empirical work also suggests that although some kind of social interactions can cause crime rates to differ in otherwise similar cities, this

is not due to the kind of multiple equilibria in the overload model. See Glaeser, Edward, Bruce Sacerdote and Jose Scheinkman, Crime and Social Interactions, NBER working paper 5026, February 1995.

5. This literature includes: John Lott, The Effect of Conviction on the Legitimate Income of Criminals, 34 Economics Letters 381 (1990); Freeman, Richard, Crime and the Employment of Disadvantaged Youth, in Adele Harrell and George Peterson, eds., Drugs, Crime and Social Isolation: Barriers to Urban Opportunity, Washington: Urban Institute Press, 1992; John Lott, An Attempt at Measuring the Total Monetary Penalty from Drug Convictions: The Importance of an Individual's Reputation, 21 Journal of Legal Studies 159 (1992); Jonathan Karpoff & John Lott, The Reputational Penalty Firms Bear from Committing Criminal Fraud, 36 Journal of Law and Economics 757 (1993); and Joel Waldfogel, Does Conviction Have a Persistent Effect on Income and Employment? 14 International Review of Law and Economics 103 (1994) .

6. John Lott, Do We Punish High-Income Criminals Too Heavily?, 30 Economic Inquiry 583 (1992).

7. Jeffrey Grogger, Arrests, Persistent Youth Joblessness, and Black-White Employment Differentials, 74 Review of Economics and Statistics 100 (1992).

8. Jeffrey Grogger, The Effect of Arrest on the Employment and Earnings of Young Men, 90 Quarterly Journal of Economics 51 (1995); Jeffrey Grogger, Criminal Opportunities, Youth Crime, and Young Men's Labor Supply, working paper, Dept. of Economics, University of California, Santa Barbara, California, February 1994.

9. See Daniel Nagin and Joel Waldfogel, The Effects of Criminality and Conviction on the Labor Market Status of Young British Offenders, 15 International Review of Law and Economics 109 (1995). Two articles that provide further empirical evidence of a wage increase, as well as theoretical explanations are: Shawn Bushway, Daniel Nagin and Lowell Taylor, The Stigmatic Impact of Criminal Records on Legitimate Employment, Working paper, Heinz School, Carnegie Mellon University, May 1995; and Daniel Nagin and Joel Waldfogel, "The Effects of Conviction on on Income Through the Life Cycle," NBER Working Paper No. 4551, November 1993.

10. The same model could be used for social stigma with appropriate changes in interpretation–for example, friends do fewer favors for those convicted of crimes, because they are revealed as less likely to reciprocate. The relationship between stigma and marriage has many of the same features, with the two added twists that (a) marriage is a pairing of women, who com-

mit fewer crimes, with men, who commit more; and (b) marriage is closer to pure matching than to a wage-mediated relationship. As a result, if all men reduce their attractiveness by engaging in crime, there may be only a small penalty in the marriage market. I conjecture that this would accentuate the multiple equilibrium problem described later in the article.

11. Actual decline is, of course, also plausible– drug and alcohol use reduce ability and concern employers quite apart from the issue of their criminality.

12. See William Dickens, Lawrence Katz, Kevin Lang, & Lawrence Summers, Employee Crime and the Monitoring Puzzle, 7 Journal of Labor Economics, 331 (1989) at 332, 335.

13. The IQ of criminal offenders is about eight points lower than that of the general population (half a standard deviation), and this does not seem to arise from measuring the IQs only of criminals who are caught. (Richard Herrnstein, Criminogenic Traits, in James Q. Wilson and Joan Petersilia, Eds. Crime, 1995, pp. 39-64 at 49.)

14. The exogeneity of α, $\underline{V}$ and $\underline{P}$ are simplifying assumptions made to highlight the effect of stigma. Quite plausibly, the reward for crime falls as the amount of crime rises, because of competition for criminal opportunities. The public penalty might either rise (from growing public concern over crime) or fall (the "overload theory" of Section I). These effects are ruled out in the present model. Note also that in this model courts do not use character evidence to stigmatize defendants, the idea behind multiple equilibria in Joel Schrag & Suzanne Scotchmer, Crime and Prejudice: The Use of Character in Evidence in Criminal Trials, 10 Journal of Law, Economics, and Organization 319 (1994).

15. Heterogeneity is imposed so that statements can be made about how the amount of crime changes with the parameters, since if individuals are identical either all are criminal or all are noncriminal. The conclusion found below that multiple equilibria can exist would remain valid even if all individuals were identical.

16. Points (a) through (d) are obvious from inspection of equation (5). Regarding point (e):

$$\frac{\partial A}{\partial \alpha} = -P - \left(\frac{1-\theta}{(1-\alpha\theta)^2} \right) y.$$

This expression is negative. Regarding point (f):

$$\frac{\partial A}{\partial \theta} = \frac{\alpha y}{1-\alpha\theta} - \frac{\alpha^2(1-\theta)y}{(1-\alpha\theta)^2} = \alpha y \left(\frac{1-\alpha}{(1-\alpha\theta)^2} \right).$$

This expression is positive under our assumption that $0 < \alpha < 1$.

17. The values of $\underline{u}$ that bound the intervals can be found by setting $\underline{A = 0}$ and θ equal to zero or to one in equation (6), yielding $\underline{u} = (V - \alpha P) - \alpha y$ and $\overline{u} = (V - \alpha P)$.

18. For a survey of studies of youth criminality, see Christy Visher & Jeffrey Roth, Participation in Criminal Careers, in Alfred Blumstein, Jacqueline Cohen, Jeffrey Roth, & Christy Visher, editors, Criminal Careers and "Career Criminals", Volume 1, (1986).

19. John Ball, Alan Ross, & Alice Simpson, Incidence and Estimated Prevalance of Recorded Delinquency in a Metropolitan Area, 29 American Sociological Review 90 (1964).

20. Robert Tillman, The Size of the 'Criminal Population': The Prevalence and Incidence of Adult Arrest, 25 Criminology 561 (1987).

21. Freeman, supra, note 2, at 172.

22. Freeman, supra, note 2, at 172. The proportions incarcerated were 3.1 and 12.7 percent.

23. Benjamin Klein & Keith Leffler, The Role of Market Forces in Assuring Contractual Performance, 89 Journal of Political Economy 615 (1981).

24. David Hirshleifer & Eric Rasmusen, Cooperation in a Repeated Prisoner's Dilemma with Ostracism, 12 Journal of Economic Behavior and Organization 87 (1989).

25. In many cases, $\underline{P = 0}$ is a reasonable approximation. The stigma arises from arrest, even if no trial follows, or conviction is followed by probation instead of imprisonment. Only 51 percent of federal and an estimated 46 percent of state felony convictions were followed by incarceration in a typical year (Federal: from 1 July 1985 to 30 June 1986, Bureau of Justice Statistics, U.S. Dept of Justice, Technical Appendix, Report to the Nation on Crime and Justice, Second Edition (1988) at 54. State: 1986 data, Bureau of Justice Statistics, U.S. Dept of Justice, Sourcebook of Criminal Justice Statistics, 1988, Table 5.31.

26. Lott, 1992, supra note 5.

27. 1993 participation rate data, from the Statistical Abstract, supra note 2, 1994, at 395, 403. Unemployment rates were 20.4 percent for age 16-19, 6.9 percent for 25-34.

28. Empirical work to determine the effects of moral decline and court

opinions is intrinsically difficult. For a survey of the work on court opinions, see Paul Cassell, Miranda's Social Costs: An Empirical Reassessment, 90 Nw. U. Law. Rev. 387 (1996).

29. Statistical Abstract, supra note 2, 1988 at 165, 1994 at 206.

30. For example, from 1961 to 1985 the arrest rate for those aged 21 to 24 rose from 8,167 to 13,054, while the rate for those aged 35 to 39 fell from 6,321 to 5,313. The pattern is even more striking for more extreme ages on each side. (Technical Appendix, supra, note 25, pp. 26-27.)

31. The effect on black males, a subpopulation easily identified by employers, may have been especially strong. The percentage of black males aged 20-24 not participating in the labor force rose from 10.2 percent in 1965 to 18.5 percent in 1971 and 21.1 percent in 1980. For white males, the figures are 14.7 percent , 16.8 percent , and 12.9 percent (from Table 8 of Murray, supra, note 2).

32. See Grogger, supra, note 8.

33. Becker, supra, note 1.

34. This idea, suggested to me by Eric Posner, is a supplement rather than a substitute for two other explanations: that multiple convictions are a more accurate sign that the convicted person was truly guilty, and that they show that he has an unusually strong tendency towards crime that requires heavier disincentives than for the average person.

35. Richard Posner writes that stigma can supplement official punishment for white-collar crime, but he misses this point, claiming instead that "The economic objection to relying on stigma for deterrence is that, like imprisonment, it is more costly to society than the pure fine (or civil penalty) because it does not yield any revenue" (Richard Posner, Optimal Sentences for White–Collar Criminals, 17 American Criminal Law Review 409 (1980) at 416). Posner is correct, however, in that fine revenue have the social benefit of replacing a certain amount of distortionary taxation.

36. There seems no constitutional objection to disclosure. In 1976, the U.S. Supreme Court established that a police department could even circulate the names of those arrested for shoplifting (even though not convicted) to local merchants (Paul v. Davis, 424 U.S. 693 (1976)). The discussion in this paragraph and the next is from Bureau of Justice Statistics, U.S. Dept of Justice, Public Access to Criminal History Record Information (1988): blotters, p. 2; dockets, p. 3; Florida, p. 19; Illinois, p. 25; and Bureau of Justice Statistics, U.S. Dept of Justice, Use and Management of Criminal History Record Information: A Comprehensive Report (1993).

37 Additional crimes requiring notification in particular states include various other offenses involving children, adultery (Arizona), bigamy (Louisiana) and voyeurism (Ohio). The best- known statute is "Megan's Law" in New Jersey. For details, see Abril Bedarf, Comment: Examining Sex Offender Community Notification Laws, 83 Calif. L. Rev. 885, 886 (1995). The trend continues as this article is being written; in May 1996, the U.S. House passed a federal version of Megan's Law by a vote of 418 to 0 (Associated Press, House passes federal version of Megan's law, Bloomington Herald-Times, May 8, 1996 at A3.

Megan's Law has been tied up in litigation for two years as of 1996, winning in the New Jersey Supreme Court (Doe v. Poritz, 142 N.J. 1, 662 A.2d 367 (1995)) and the U.S. Third Circuit (Artway v. Attorney General, 81 F.3d 1235 (3d Cir. 1996)), but delayed by preliminary injunctions.

President Clinton did sign the federal bill, but stigma may become an issue in the 1996 elections nonetheless. Senator Dole has called for disclosing juvenile criminal records to schools, courts, and employers, but President Clinton has yet to take a stand on this. Dole Seeks to Get Tough on Young Criminals, Los Angeles Times, Sunday, July 7, 1996 at A16. For current law, see Neal Miller, Bureau of Justice Statistics, U.S. Dept of Justice, State Laws on Prosecutors' and Judges' Use of Juvenile Records (1995).

38. Bedard, supra note 37, at 904, 905.

39. Discrimination against criminals is generally legal, but it has become entangled in racial discrimination law. In one case, a plaintiff was refused employment because of his 14 arrests. Judge Hill said: "There is no evidence to support a claim that persons who have suffered no criminal convictions but have been arrested on a number of occasions can be expected, when employed, to perform less efficiently or less honestly than other employees. In fact, the evidence in the case was overwhelmingly to the contrary. Thus, information concerning a prospective employee's record of arrests without convictions is irrelevant to his suitability or qualification for employment." Gregory v. Litton Systems, Inc., 316 F. Supp. 401 (C.D. Cal 1970), affirmed 472 F 2.d 631 (9th Cir. 1972).

40. Note, however, that the opportunity cost of crime would fall for workers who had not been criminal in the past, since they would receive the same pooled wage as the criminal workers.

41. A subtly different argument with similar implications is that by raising the criminal's income, legitimate employment reduces his marginal utility of income and his temptation to commit property crimes. See Eric Rasmusen, An Income-Satiation Model of Efficiency Wages, 30 Economic

Inquiry 467 (1992).

42. Ironically enough, this example of privatization and information disclosure has been attacked by the usually pro-market *Wall Street Journal* (editorial, Flawed Law, July 9, 1996): "This law offloads the problem onto the public itself. At best it's an incentive to vigilantism. At worst it will extinguish the value of homes; who in their right mind would move into a known Megan's Law neighborhood?" Allowing people to avoid the hazards of employing or living near offenders is, of course, precisely the point. Wages and housing values fall for some people, but rise for others, and rise more, since better information increases efficiency.

FIGURES Figure 1: Multiple Equilibria

Figure 2: Shifting Equilibria When Criminal Penalties are Reduced