-- AND M. MELEU [2000], "Enforcement of Contracts with Adverse Selection in LDCs," Working Paper, Atelier de Recherche Quantitative Appliquée aux Développement Economique, University of Toulouse I, Toulouse, France.
-- AND T. N'GUESSAN [2000], "Group Lending with Adverse Selection," *European Economic Review*, 44, 773–784.
-- AND P. REY [2000], "Collusion and Group Lending with Moral Hazard," Working Paper, Institut d'Economie Industrielle, University of Toulouse I, Toulouse, France.
RAI, A., AND T. SJÖSTRÖM [2000], "Efficient Lending in Village Economies," Working Paper, Kennedy School of Government, Harvard University, Cambridge, MA.
STIGLITZ, J. [1990], "Peer Monitoring and Credit Markets," *World Bank Economic Review*, 4, 351–366.
VARIAN, H. [1990], "Monitoring Agents with Other Agents," *Journal of Institutional and Theoretical Economics*, 146, 153–174.

Jean-Jacques Laffont
Institut d'Economie Industrielle
Université des Sciences Sociales de Toulouse
Place Anatole France
31042 Toulouse Cedex
France
E-mail: laffont@cict.fr

Tchétché N'Guessan
CIRES
Université de Cocody
Abidjan
Côte d'Ivoire

An Economic Theory of Deviance

by

JEONG-YOO KIM AND GANG LEE*

We develop a model of deviance by incorporating the labeling effect into rational choice theory. In our model, we provide an explanation of the process through which a deviant is stigmatized and explore theoretically the relationship between the experience of having deviated and the incentive for deviation. Surprisingly, our study finds that an ex-deviant is not necessarily more likely to deviate, contrary to widely held belief. This is because the ex-deviant may hesitate to violate the norm one more time for fear of being labeled as a pathological deviant. (JEL: K 49)

1 Introduction

The term "deviance" may be outside the technical vocabulary of economists, but it is commonly used by sociologists. It is defined by the latter as a violation of a social norm (that is, a standard code) about how human beings ought or ought not to act under given circumstances.[1]

Deviant behavior usually evokes formal and informal punishment, restrictions, or other controls of society. These controls constrain most people to conform to social norms. Despite the social sanctioning and controlling, however, we sometimes observe deviant behavior around us. Why do some people engage in such deviant behavior even if social punishments are expected? Sociologists have attempted to explain this in various ways. One group of sociologists explain deviance in terms of broad social conditions in which it is most likely to grow, by looking at the structural characteristics of society and groups within society (e.g.,

* We would like to thank two anonymous referees, David George, Shoshana Grossbard-Shechtman, and Louis Levy-Garboua for helpful comments, as well as participants in the seminar of the Society for the Advancement of Behavioral Economics (SABE) held at San Diego State University, in the conferences of the Korean Sociology Association, and in seminars at Columbia University, Stanford University, and Kyunghee University. We also wish to acknowledge the financial support from the nondirected research fund of the Korea Research Foundation.

[1] Criminologists usually distinguish *crime* from deviance that is against nonbinding social norms but that is not actually a violation of criminal law. Noncriminal deviance includes religious fanaticism, certain political belief systems, certain dressing and hairdressing styles, excessive drinking and gambling, hooliganism, mental disorders, etc. However, the difference is not essential from the economic perspective.

Journal of Institutional and Theoretical Economics
JITE 157 (2001), 499–519 © 2001 Mohr Siebeck – ISSN 0932-4569

MERTON [1938], CLOWARD AND OHLIN [1960]). Another group explains it in terms of the characteristics of individuals, focusing on those characteristics that are most highly associated with learning deviant acts (e.g., SUTHERLAND [1947]). Yet another line of thought stresses the importance of the labeling (or stigmatizing) effect. According to this, interpersonal reactions to deviance may significantly increase the likelihood of subsequent deviant behavior (e.g., H. BECKER [1963]). In contrast with the first two schools of thought, which are mainly interested in primary deviance, this theory focuses on secondary deviance.[2]

However, rational-choice theorists, including such economists as Gary Becker, have argued that an individual chooses to conform to or violate the social norm by calculating the risk of social punishment or pain and weighing it against the potential gain and pleasure derived from the act. They thus explain that an individual engages in deviance because the benefit from violating the norm exceeds the cost of it.

Rational-choice theory predicts that individuals in different conditions (in time or space) will have different incentives to deviate. For example, individuals living in poor neighborhoods are more inclined to commit theft, robbery, rape, etc.,[3] and the crime rate is higher in a recession than in a boom.[4] Also, the theory suggests that the incentive for an individual to deviate may vary with changing conditions. One of the other important factors affecting an individual's incentive to deviate must be his experience of violating norms and resultant perceptions of his mentality by other community members. According to a special report published by the U.S. Department of Justice under the Federal Justice Statistics program, of the 33,855 offenders returning to prison between 1986 and 1997, 54% returned within 1 year of being released, an additional 34% returned within 2 years, and 12% returned after 2 to 3 years.[5] These statistics seem to constitute strong empirical evidence for the correlation between the number of experiences in deviant behavior and the incentive to deviate. Nonetheless, to the best of our limited knowledge, there has been no theoretical study exploring the relationship between experience in deviation and the incentive to deviate.

To establish a theoretical foundation for the relation between them may have important policy implications. If a deviant act by someone can cause harm to others, as in the case of crime, the public authority needs to monitor *ex ante* individuals who are likely to violate norms in order to deter deviance, and can minimize monitoring costs by monitoring intensively individuals who are most likely to deviate. Also, if the relation is caused by the labeling of a deviant, the public authority will be able to reduce the deviance rate by taking positive policy measures to mitigate the labeling effect.

[2] In sociologists' terminology, secondary deviation refers to the responses that people make to problems created by the societal reaction to their deviance, while primary deviation refers to the enactment of deviant behavior itself.

[3] A study found that two-thirds of the homicides in Cleveland occurred in 12% of the city, primarily in black, inner-city areas (BENSING AND SCHROEDER [1960]). Similar patterns have been documented in other large cities in the U. S. as well as other countries (REISS JR. AND ROTH [1993]).

[4] COOK AND ZARKIN [1984] find that there is an increase in the number of robberies and burglaries during recessions.

[5] See the special report "Offenders Returning to Federal Prison, 1986–1997", Bureau of Justice Statistics, Department of Justice (September 2000).

In this paper, we develop a theory of deviance by incorporating the labeling effect into BECKER's [1968] framework, that is, by combining the labeling theory with rational-choice theory.[6] According to our model, a rational person may find it in his interest to follow the social custom even if his short-run benefit from breaking it exceeds the cost of doing so, if he takes into account the long-run social interaction. By considering this general model, we will be able to tackle two important issues: through what process a deviant is labeled, and, more importantly, whether or not a deviant has more incentive to violate the norm again after further experiences of violation.

The crucial feature of this model that differentiates it from BECKER [1968] is that we explicitly incorporate into Becker's model the social sanction of labeling behavior as deviant. In general, two ways can be conceived of for introducing social sanctions explicitly into the model. One is to assume it exogenously (e.g., AKERLOF [1980], COLEMAN [1987], GLAESER, SACERDOTE, AND SCHEINKMAN [1996]), and the other is to impose it endogenously (e.g., BERNHEIM [1994], AKERLOF [1997]). In AKERLOF [1980], deviance affects the deviant's utility directly by damaging his reputation, and in BERNHEIM [1994], it affects the deviant's utility indirectly through affecting the beliefs of others. In this paper, we follow the framework of BERNHEIM [1994] in that we decompose utility into intrinsic utility derived directly from the deviant's action and extrinsic utility derived from others' perceptions, and we suppose that a deviant is sanctioned endogenously by affecting the public's beliefs about his mentality. Our approach differs from BERNHEIM's [1994] in that he focuses mainly on conforming behavior, whereas we focus on deviant behavior. Also, in BERNHEIM [1994], the intrinsic gain from departing from social norms is fixed for any individual, so that an individual who has deviated once will keep doing so, since he has no reputation to lose. This seems unrealistic in view of the fact that each individual deviates at some times and does not at other times. Deviance is a phenomenon contingent on the situation, not a phenomenon structurally built into someone.

Sociologists who believe in the labeling theory of deviance (e.g., BRAITHWAITE [1989]) assert that stigmatizing a person as delinquent or criminal more often encourages than discourages subsequent deviant behavior and that the stigmatization of deviants puts them at high risk of behaving according to the label, playing out the role of a deviant, and developing self-concepts as irrevocably deviant. However, this paper will highlight the possible positive aspect of the stigma effect, rather than the negative effect most sociological literature focuses on. The main finding is that there is some range of the public's perception within which an individual who has experi-

[6] RASMUSEN [1996] also considers the stigma effect within the framework of rational-choice theory. However, he does not address the issue of secondary deviance, which is the essence of both our paper and the labeling theory.

enced violating norms once is less likely to redeviate than an individual who has no such experience. To put it another way, an ex-deviant is not necessarily more likely to deviate, contrary to widely held belief. This is because the ex-deviant may hesitate to violate the norm one more time lest he should be labeled as a pathological deviant. That is, the possibility of severe stigmatization has a deterrent effect.[7] To illustrate, suppose an economist keeps protesting against journal editors' editorial decisions. They may at first consider his claim seriously, but if the fact of his repeated protests is eventually known to the academic circle, other economists will be negatively impressed by him. Therefore, knowing this, he will refrain from protesting even if he does not agree with editors' decisions at some stage.

Of course, it is often observed that a feeble-minded and naive juvenile suddenly becomes a recidivist once he commits a crime. Whether an experience of deviance reinforces or weakens the incentive of an individual to redeviate will depend on many factors, including the degree of others' repugnance against the deviant behavior, the possibility that the deviant behavior was made by mistake, the possibility that it was inevitable, etc. If a certain deviant behavior is considered not serious, so that its labeling effect is mild, the one-time deviant will be rather reluctant to redeviate, since another deviance might cause a serious stigmatization. However, as he repeatedly perpetrates deviant behavior, his reputation becomes seriously damaged, and eventually he will not hesitate to redeviate when he is labeled a loathsome anomaly.

No claim similar to our main result is found in the literature except in the context of divorces. For example, CHERLIN [1978, 641] noticed that "The previously divorced may be more hesitant to divorce again because of the stigma attached to divorcing twice." However, it is beyond doubt that the observation of this pattern of behavior will not be limited to divorce cases but will be universal. This paper adds value to the existing literature in that it contains a formal analysis of an abstract, general model conveying a behavioral principle.

The organization of this paper is as follows: We set up a static model in section 2. In section 3, we make a comparative-statics analysis and derive the main results. In section 4, we relax some restrictive assumptions made in section 2 and provide a stochastic dynamic model. Section 5 contains some discussions, and section 6 draws some policy implications. In section 7, we contrast unofficial stigmatization with official punishment. Concluding remarks and caveats follow in section 8.

2 *The Model*

We consider a simple model explaining deviant behavior. The population consists of two types of individuals, rational and pathological. A rational individual, by

[7] In the economic literature, RASMUSEN [1996] recognized this possibility earlier, although he defined stigmatization differently.

definition, makes a decision by maximizing his expected utility, reflecting the cost and benefit of the deviant behavior. On the other hand, a pathological individual does not maximize expected utility, but behaves impulsively. Thus we define a pathological individual as one who always engages in deviant behavior. We assume that the proportion of pathological individuals in the population is $\lambda \in (0, 1)$ and that it is common knowledge.[8] This model can be interpreted alternatively as having a single individual who can be either rational or pathological. In this interpretation, λ can be viewed as the prior belief that the individual is pathological.

A rational individual's utility function has two components. One is the intrinsic utility derived directly from a behavior, and the other is the extrinsic utility derived from social interaction with other community members. Since an individual lives with others in his community, he must take into account the consequences of his decision in social exchange.

If he disobeys social customs and commits a deviant act, it entails direct benefits and costs to him.[9] Let the net benefit from the deviant behavior be denoted by v.[10] Then v is assumed to have the nonatomic distribution function $F(v)$ over V with the corresponding density function $f(v)$, where $0 \in V \equiv [\underline{v}, \bar{v}]$. The value of v is private information of the individual. If an individual is believed to be pathological by the whole population, it incurs disutility w to him.[11] w is assumed to be constant.[12] Then a rational individual's utility can be written as $U = U_I + U_E$, where U_I, U_E are symbols for intrinsic and extrinsic utility respectively. We assume that $U_I = v$ if he deviates from social norms, and we normalize the intrinsic utility when he conforms to social norms to $U_I = 0$. The public updates its belief that an individual is pathological, based on its observation of his behavior. Let $\hat{\lambda}$ be the posterior belief. Then his utility is $U = v - \hat{\lambda}w$ if he deviates, and $U = -\lambda w$ (since $\hat{\lambda} = \lambda$) if he conforms to the social norms. The extrinsic utility $-\lambda w$ when an individual follows norms reflects the externality generated from the presence of pathological people in the population. Here, $\hat{\lambda}$ is the common belief by the public and will be called the *reputation* of an individual. If $\hat{\lambda}$ is higher (lower, respective-

[8] Some deviant acts are clearly unplanned, as WITTE AND TAUCHEN [1994] noticed. This implies that some proportion of deviance cannot be explained in the absence of non-utility-maximizers.

[9] In the divorce case, for example, direct benefits of divorcing include freedom from a painful marriage, relief from financial obligations and responsibilities for his family, etc., while direct costs include loneliness.

[10] In this paper, externalities generated by deviant behavior are not considered explicitly; neither is the social welfare function. First, that is because deviant behavior, in most cases, has only a marginal effect on others' utility. Deviant behavior other than crimes usually makes others feel unpleasant, but does not hurt them seriously. Second, it is because our main interest is in the individual behavior rather than in social welfare.

[11] If a person is stigmatized as abnormal, people will shun him and he will lose a lot of opportunities to work, since firms will not hire him even though they know that he has high productivity. For empirical evidence of the stigma effect, see RASMUSEN [1996].

[12] One can imagine an implicit subgame describing social interaction following deviant behavior. w can be thought as the equilibrium value endogenously determined in the subgame.

ly), it will be said that he has a *worse* (*better*, respectively) reputation. We are implicitly assuming that the stronger the belief that a given individual is pathological, the larger the proportion of the population that keeps away from him. So the disutility a deviant must suffer from the severance of social relations will be proportional to the posterior belief. For expository convenience, we are simply assuming that the expected disutility of violating norms due to social interaction, when the resulting posterior belief is $\hat{\lambda}$, is $\hat{\lambda}w$.

Suppose a given individual is a rational person attaching no intrinsic value to social interaction, i.e., $w = 0$. Then all that matters to him is his intrinsic utility. So he will conform to social customs if his intrinsic utility when conforming to social customs ($U_I = 0$) exceeds the intrinsic utility when breaking them ($U_I = v$), i.e., $v < 0$; he will deviate if $v > 0$; and he will randomize if $v = 0$. As a tie-breaking rule, we assume that he deviates with probability 1 if $v = 0$. Then the necessary and sufficient condition for a rational individual to commit an act of deviant behavior in this case is that $v \geq 0$.

Now, let us consider how the decision-making of a rational individual is affected by considering the extrinsic utility from social interaction. If he conforms, he gets $-\lambda w$. If he deviates, the public updates the posterior belief that he is pathological according to the Bayes's law, i.e., $\hat{\lambda} = \dfrac{\lambda}{\lambda + (1-\lambda)(1-F(v^*))}$, where v^* is the net benefit from deviation that makes him indifferent between deviating and conforming. Thus, the utility when he deviates is $v - \dfrac{\lambda}{\lambda + (1-\lambda)(1-F(v^*))} w$. Then v^* can be determined by

$$-\lambda w = v^* - \frac{\lambda}{\lambda + (1-\lambda)(1-F(v^*))} w. \tag{1}$$

Rearranging (1) yields

$$v^* = \left\{ \frac{\lambda}{\lambda + (1-\lambda)(1-F(v^*))} - \lambda \right\} w \equiv \phi(v^*). \tag{2}$$

We have $\phi(\underline{v}) = 0$ and $\phi(\bar{v}) = (1-\lambda)w$. Thus, the necessary and sufficient condition for the existence of $v^* \in \text{Int } V \equiv (\underline{v}, \bar{v})$ such that $v^* = \phi(v^*)$ is

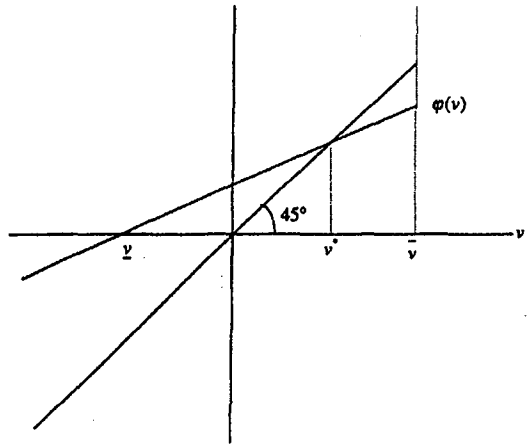
$$(1-\lambda)w < \bar{v}. \tag{NS}$$

Furthermore, if we assume that $\phi'(v) \leq 0$ for all $v \in V$, then he deviates if $v \geq v^*$ and conforms if $v < v^*$, where $v^* \in (0, \bar{v})$, since $\phi'(v) > 0$. (See figure 1.) If $(1-\lambda)w = \bar{v}$, a deviant behavior may occur, but only with zero probability, since $v^* = \bar{v}$ and $F(v)$ has no atom. Also, if $(1-\lambda)w > \bar{v}$, no $v \in V$ satisfies (2). This implies that there exists no $v \in V$ making the individual indifferent between following the norm and breaking it if the posterior belief $\hat{\lambda}$ is formed according to Bayes's law. This in turn implies that the individual either strictly prefers conforming to the norm or strictly prefers deviating from it for all v.

First, consider the possibility that an individual strictly prefers deviating for all v. If that were the case, the posterior belief would be $\hat{\lambda} = \lambda$. Then, even if an individual violated the norm, there would be no stigma effect. Therefore, an individual with $v < 0$ would conform, which is contradictory. Now, suppose that an individual strictly prefers conforming to the norm for all v in equilibrium. Then, when an individual deviates from it, he is believed to be pathological with posterior probability $\hat{\lambda} = 1$. Since $-\hat{\lambda}w = -w < \lambda w \ \forall v$, it can be an equilibrium that a rational individual conforms to social norms for all v. This can occur especially when w is sufficiently large that a departure from the social norm seriously impairs the reputation of the deviant. The following proposition summarizes this.

Figure 1

Determination of v^*



Proposition 1: Assume that $\phi''(v) \leq 0$ for all $v \in V$. Then the necessary and sufficient condition for a rational individual to exhibit deviant behavior with positive probability is that $(1-\lambda)w < \bar{v}$ (NS), i.e., $\exists v^* \in (0, \bar{v}) \ni v^* = \phi(v^*)$. Furthermore, in this case, (i) he deviates if $v \geq v^*$ and conforms if $v < v^*$, and (ii) $v^* > 0$.

If $v \geq v^*$, even a rational individual deviates at the expense of higher probability of being considered as pathological, since the cost of conforming to social norms is too high. Also, (ii) implies that a rational individual with private information $0 \leq v < v^*$ does not deviate even if his direct benefit from deviation exceeds the direct cost, lest he should lose his future benefit from social interaction.

If condition (NS) is satisfied, we have $0 < F(v^*) < 1$, so that $\lambda < \hat{\lambda} < 1$. If this condition is violated, it must be that $\hat{\lambda} = 1$.[13] Thus, in either case, we have $\hat{\lambda} > \lambda$.

Proposition 2: Given $\lambda \in (0, 1)$, one has $\hat{\lambda} > \lambda$.

[13] If $(1-\lambda)w = \bar{v}$, it is obvious that $F(v^*) = 1$, so that $\hat{\lambda} = 1$.

This proposition says that a deviant behavior strictly increases the posterior belief that the deviant is abnormal. If $\hat{\lambda}=\lambda$, a deviant behavior will not entail any stigma effect at all, so that a rational person's decision will be identical to the myopic decision that did not take into account the future payoff. This proposition implies that a long-sighted decision by a rational person must differ from the myopic decision.

3 *Comparative Statics*

Assuming that condition (NS) is satisfied, we will perform a comparative-statics analysis. The following propositions describe the comparative-statics properties of the equilibrium outcome.

Proposition 3: $v^*=0$ if $w=0$, and v^* is increasing in w.

Proof: Obvious from figure 1. *Q.E.D.*

This proposition says that deviance is less likely to occur when the benefit from social interaction is larger. For the likelihood of deviant behavior is $\lambda+(1-\lambda)[1-F(v^*)]$.

Corollary 1: $\hat{\lambda}$ is increasing in w.

Proof: $\dfrac{\partial\hat{\lambda}}{\partial w}=\dfrac{\partial\hat{\lambda}}{\partial v^*}\dfrac{\partial v^*}{\partial w}>0$, since $\dfrac{\partial\hat{\lambda}}{\partial v^*}>0$. *Q.E.D.*

This corollary implies that a larger w makes an individual less inclined to violate the norm, so that one act of deviation hurts his reputation more seriously.

Proposition 4: (i) $\dfrac{\partial v^*}{\partial\lambda}>0$, or equivalently, $\dfrac{\partial\hat{\lambda}}{\partial\lambda}>1$, for all $\lambda\in(0,\lambda_\varepsilon)$ for some $\lambda_\varepsilon>0$.[14] (ii) $\dfrac{\partial v^*}{\partial\lambda}<0$, or equivalently, $\dfrac{\partial\hat{\lambda}}{\partial\lambda}<1$, for all $\lambda\in(\frac{1}{2},1)$.

Proof: See the appendix.

In an intermediate range of λ ($\lambda_\varepsilon<\lambda<\frac{1}{2}$), however, the direction of a change in v^* with an increase in λ is ambiguous.

One implication of this result can be given in the following way. Deviant behavior of an individual makes people believe with higher probability that he is pathological, i.e., $\hat{\lambda}>\lambda$ by proposition 2. Then, after the deviation, $\hat{\lambda}$ becomes the new prior belief of the public that he is pathological, until he perpetrates another deviant behavior. If $\hat{\lambda}$ is very small, an ex-deviant whose reputation is increased to $\hat{\lambda}$ is

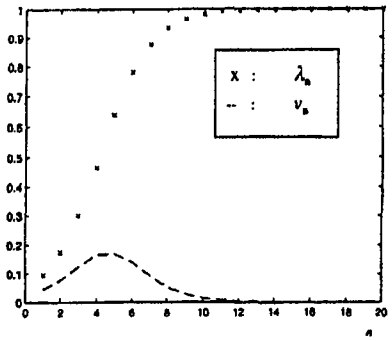
[14] This result is in contrast with Rasmusen's assertion, "The stigma from a first conviction is greater than from subsequent convictions, and after enough convictions the marginal effect is negligible" (RASMUSEN [1996, 536]).

less likely to deviate. This is mainly because another deviation would greatly increase the probability that he is believed to be pathological. However, if $\hat{\lambda}$ is large (more specifically, $\hat{\lambda}>\frac{1}{2}$), we can be sure that the ex-deviant is more likely to deviate. The intuitive reason is that, if $\hat{\lambda}$ is large, the belief that he is pathological is not increased very much even if he deviates one more time. The upshot is that an ex-deviant is not necessarily more likely to deviate again. If he is, for example, deviating for the first time, he will be less inclined to commit repeat deviance, in order not to be believed to be habitual. On the contrary, if he has already deviated many times, and thus the belief that he is pathological is high enough, he will be more inclined to deviate again, since he does not have much reputation to lose.

Figures 2 to 7 provide simulated examples of the equilibrium dynamics of an individual's behavior pattern. For the numerical simulation, it is assumed that v has a uniform distribution that is symmetric about 0, i.e., $\underline{v}=-\bar{v}$ and $F(v)=\dfrac{v+\bar{v}}{2\bar{v}}$. Here v_n and λ_n are defined recursively as satisfying $v_n=(\lambda_n-\lambda_{n-1})w$ and $\lambda_n=\dfrac{\lambda_{n-1}}{\lambda_{n-1}+(1-\lambda_{n-1})(1-F(v_n))}$, where $\lambda_0=\lambda$. That is, v_n represents a marginal individual who has deviated $n-1$ times, and λ_n is the posterior belief of the public when it has observed an individual deviating n times. The monotonicity of the lines marked × shows that the posterior belief becomes higher as the number of deviations is increased, and the nonmonotonicity of the broken lines indicates that the incentive to deviate is decreasing at the early stage of repeat deviations and then increases.

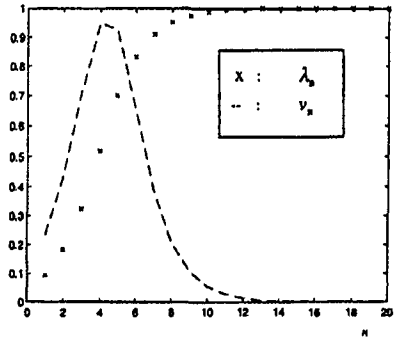
A natural question is to what extent this conclusion is applicable. For which among various kinds of deviant behavior can it provide a useful prediction? That may depend on how much a one-time deviation will affect the public's perception. If it changes the public's perception to $\hat{\lambda}\in(0,\lambda_\varepsilon)$, the deviant will become more hesitant to engage in another deviation. On the other hand, if it changes the perception to $\hat{\lambda}\in(\frac{1}{2},1)$, the stigma effect will strengthen his incentive to deviate. Various factors must be considered to tell under which case a specific type of deviant behavior falls, but the size of w is obviously one of the key determinants. According to corollary 1, an act of deviance increases $\hat{\lambda}$ more if w is larger, where w can be interpreted as an index of others' repugnance against a particular deviant behavior. This implies that the seriousness of a deviant behavior or the degree of others' repugnance against it can be a good criterion for the applicability of our result. For example, it is quite rare to observe a famous actor or actress involved in a shameful scandal more than once during a short period of time, unless he or she is already notorious for his or her habitual misconduct. In this type of deviance, w is considered to be relatively small. For ordinary people are not afraid of interacting with such a person, since they do not expect any harm to be inflicted by that type of deviant behavior. On the other hand, in the case of a felony like murder or robbery, almost anybody could be a victim of the deviant behavior, implying that w is quite large. Consequently, one-time deviation in this case would make $\hat{\lambda}$ close to 1, so that the deviant is stigmatized.

Figure 2



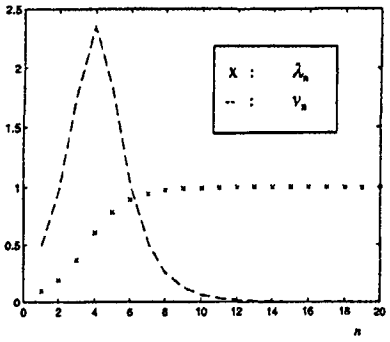
Parameter values: $\lambda_0 = 0.05$, $w = 1$, $\bar{v} = 10$

Figure 3



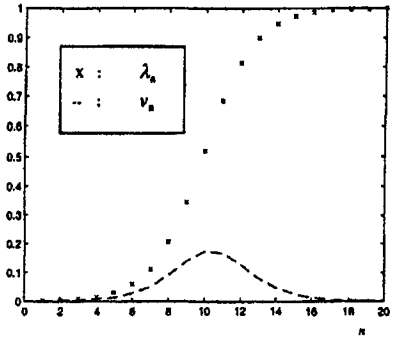
Parameter values: $\lambda_0 = 0.05$, $w = 5$, $\bar{v} = 10$

Figure 4



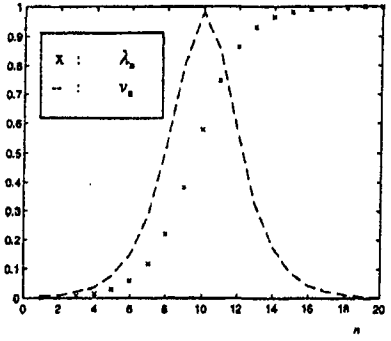
Parameter values: $\lambda_0 = 0.05$, $w = 10$, $\bar{v} = 10$

Figure 5



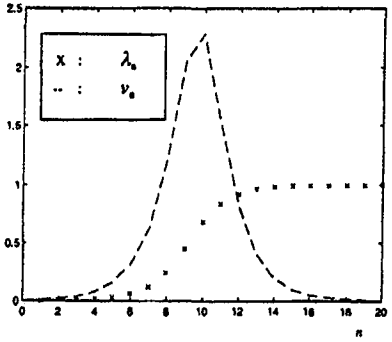
Parameter values: $\lambda_0 = 0.001$, $w = 1$, $\bar{v} = 10$

Figure 6



Parameter values: $\lambda_0 = 0.001$, $w = 5$, $\bar{v} = 10$

Figure 7



Parameter values: $\lambda_0 = 0.001$, $w = 10$, $\bar{v} = 10$

4 Stochastic Dynamic Model

The model provided in section 2 is restrictive in two respects. First, the rationality of rational individuals was bounded in the sense that they did not take into account the possibility that they would redeviate in the future, when they contemplate committing an act of deviant behavior. Second, an individual's reputation was updated only when he had engaged in deviant behavior. An act in accordance with the norms did not improve his reputation. This feature comes from the implicit assumption that conforming behavior was not observable to the public. Thus the public was not given a chance to update its posterior belief when an individual obeyed social norms.

In this section, we modify the restrictive features by assuming that individuals are forward-looking and that both deviant behavior and conforming behavior are observable. Also, to make things realistic and interesting, we introduce uncertainty about deviants' decisions into the model.[15]

We consider the following dynamic model. An individual lives for two periods. The first period can be thought of as his adolescence, and the second period as his senescence. The chances of getting short-run gains from violating the norm arrive stochastically in each period. The chances are given to an individual independently in each period and the probability that a chance arrives in each period is denoted by p. Also, net benefits from breaking the norm are known to the individual at the same time that the random events arrive. Let v_t denote the net benefit of deviating from the norm in period t. Then the v_t's are independently and identically distributed according to $f(v)$ over V. Values of v_t's are private information of the decision maker, while $f(v)$ is common knowledge. The action of an individual at some period affects not only his (per-period) intrinsic utility, but also his (per-period) extrinsic utility as in section 2. Second-period utility is discounted by the discount factor $\delta\,(< 1)$. The public's prior belief that an individual is pathological is λ_0, and it updates the posterior belief according to Bayes's law, based on its observation of the individual's action in each period.

Let $d \in D \equiv \{0, 1\}$ be an action (or a decision) of the individual, where $d = 1$ denotes deviating and $d = 0$ denotes conforming to the norm. Also, let $\lambda_t(d_t; \lambda_{t-1})$ be the public's period-t posterior belief when it observes an action d_t in period t and the period-$(t-1)$ belief is λ_{t-1}. Then $\lambda_t(d_t; \lambda_{t-1})$ is given by

$$\lambda_t(0;\lambda_{t-1}) = \frac{(1-p)\,\lambda_{t-1}}{1-p+p(1-\lambda_{t-1})\,F(v_t^*)} \equiv g(\lambda_{t-1}, v_t^*, p), \tag{3}$$

$$\lambda_t(1;\lambda_{t-1}) = \frac{\lambda_{t-1}}{\lambda_{t-1}+(1-\lambda_{t-1})\,(1-F(v_t^*))} \equiv h(\lambda_{t-1}, v_t^*, p), \tag{4}$$

[15] If we allow for the observability of nondeviant behavior without incorporating uncertainty, the analysis will become unrealistic and trivial, since the observation that an individual follows norms in a certain period makes the public believe that he will never be pathological for the rest of his life, even if he engages in deviant behavior in the very next period.

where v_t^* is the type of the individual who is indifferent between $d=0$ and $d=1$ in period t.[16] Notice that $h(\lambda, v, p)$ does not depend on p and that $g(\lambda, v, p) < \lambda < h(\lambda, v, p)$ for any $v > \underline{v}$, for any $p \neq 0$, and for any $\lambda \in (0, 1)$. Also notice that if $\lambda = 1$, we have $g(1, v, p) = h(1, v, p) = 1$ for any v and p.[17] We will call this belief *complete stigma*, and an individual with the reputation $\lambda = 1$ will be said to be *completely stigmatized*. If an individual is completely stigmatized, his reputation cannot be recovered at all, no matter what he does in the successive period.

The following lemmas will be useful to characterize our equilibrium.

Lemma 1: $g(\lambda, v, p)$ is strictly decreasing in p, $\forall p \in (0, 1)$, $\forall v > \underline{v}$, $\forall \lambda \in (0, 1)$, and $g(\lambda, v, 0) = \lambda$.

Proof: Obvious from equation (3). *Q.E.D.*

The intuition is crystal clear. An individual's conforming behavior in a situation with a higher chance of getting short-run benefits by deviation should imply more likelihood that he is rational.

Lemma 2: (i) $h(\lambda, v, p)$ is strictly increasing in v, $\forall \lambda \in (0, 1)$. (ii) $g(\lambda, v, p)$ is strictly decreasing in v, $\forall p, \lambda \in (0, 1)$. (iii) Accordingly, $h(\lambda, v, p) - g(\lambda, v, p)$ is strictly increasing in v, $\forall p \in [0, 1]$, $\forall \lambda \in (0, 1)$.

Proof: Straightforward. *Q.E.D.*

To go on with our analysis, we introduce some notation for various equilibrium continuation payoffs of a rational individual given his reputation λ in period t:

$U^t(\lambda)$ = the *ex ante* expected payoff before a random event is realized,

$U_Y^t(v, \lambda)$ = the expected payoff when the chance for a decision is given and the short-run gain from deviation is v,

$U_N^t(\lambda)$ = the expected payoff when the chance is not given,

$U_0^t(v, \lambda)$ = the payoff that a rational individual is expected to obtain by conforming to the norm when the chance to make a decision is given and the short-run gain from deviating is v,

$U_1^t(v, \lambda)$ = the payoff that a rational individual is expected to obtain by deviating when the chance to make a decision is given and the short-run gain from deviating is v.

[16] Bayes's law says that $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. Let A and B respectively be the event that a person is pathological and the event that he conforms in period t. Then $P(A) = \lambda_{t-1}$, $P(B \mid A) = 1$, and $P(B) = 1 - p + p(1 - \lambda_{t-1}) F(v_t^*)$, so that the derivation of $\lambda_t(0; \lambda_{t-1})$ is immediate. Similarly, $\lambda_t(1; \lambda_{t-1}) = P(A \mid B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{\lambda_{t-1}}{1 - (1 - \lambda_{t-1}) F(v_t^*)}$, where B^c is the complement of B.

[17] In the language of stochastic processes, $\lambda_t = 1$ is an absorbing state.

Then it is not difficult to see that

$$U^t(\lambda) = p \int_{\underline{v}}^{\bar{v}} U_Y^t(v, \lambda) f(v)\, dv + (1 - p) U_N^t(\lambda), \tag{5}$$

$$U_Y^t(v, \lambda) = \max\{U_1^t(v, \lambda), U_0^t(v, \lambda)\}, \tag{6}$$

$$U_1^t(v, \lambda) = \begin{cases} v - h(\lambda) w + \delta U^2(h(\lambda)) & \text{if } t = 1, \\ v - h(\lambda) w & \text{if } t = 2, \end{cases} \tag{7}$$

$$U_0^t(v, \lambda) = U_N^t(\lambda) = \begin{cases} -g(\lambda) w + \delta U^2(g(\lambda)) & \text{if } t = 1, \\ -g(\lambda) w & \text{if } t = 2. \end{cases} \tag{8}$$

We will be interested only in the equilibrium in which a rational individual deviates from norms with positive probability in both periods. This implies that $v_1^*, v_2^* \in \text{Int } V$, and thus, v_t^*, $t = 1, 2$, are determined by

$$U_0^t(v_t^*, \lambda_{t-1}) = U_1^t(v_t^*, \lambda_{t-1}), \quad t = 1, 2. \tag{9}$$

First, let us consider the decision of a rational individual in the second period. $v_2^*(\lambda_1)$ must satisfy

$$v_2^*(\lambda_1) = \varphi(v_2^*(\lambda_1), \lambda_1, p), \tag{10}$$

where $\varphi(v_2, \lambda_1, p) = \{h(\lambda_1, v_2, p) - g(\lambda_1, v_2, p)\} w$. Let φ_i and φ_{ii} be the first and the second derivative of φ with respect to the ith argument. Then, assuming that $\varphi_{11}(v_2, \lambda_1, p) \leq 0$ for all v_2, λ_1, p, we have the following propositions.

Proposition 5: There exists a unique $v_2^*(\lambda_1) \in (0, \bar{v})$ if and only if $w < \frac{1 - p\lambda}{1 - \lambda} \bar{v}$. In this case, the individual deviates in the second period if $v_2 \geq v_2^*(\lambda_1)$ and conforms if $v_2 < v_2^*(\lambda_1)$.

Proof: The result follows from the facts that $\varphi(\underline{v}_2, \lambda_1, p) = 0$, φ is continuous in v_2 and $\varphi_1(v_2, \lambda_1, p) > 0$, $\forall p \in [0, 1]$, $\forall v_2$. (See figure 8.) *Q.E.D.*

Proposition 6: $v_2^*(\lambda_1)$ is strictly increasing in p for all $\lambda_1 \in (0, 1)$.

Proof: Since $\varphi_3 > 0$, the result is immediate from figure 9. *Q.E.D.*

The intuition behind this proposition is that, given the reputation, an individual has more incentive to conform to the norm in the second period as the event probability becomes higher, since a higher probability of random events will make his reputation obtained as a result of conforming better according to lemma 1.

To see the decision of a rational individual in the first period, the following lemmas are helpful.

Lemma 3: $U^2(\lambda)$ is strictly decreasing in λ.

Proof: This is immediate, since both of $U_1^2(v, \lambda)$ and $U_0^2(v, \lambda)$ are strictly decreasing in λ. *Q.E.D.*

Figure 8

Determination of $v_2^*(\lambda_1)$

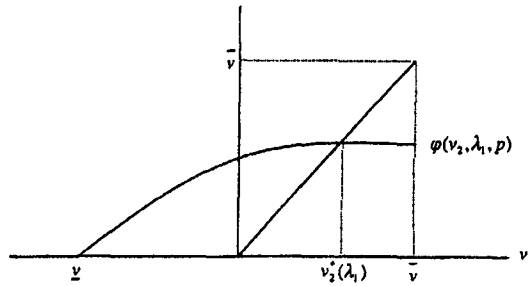
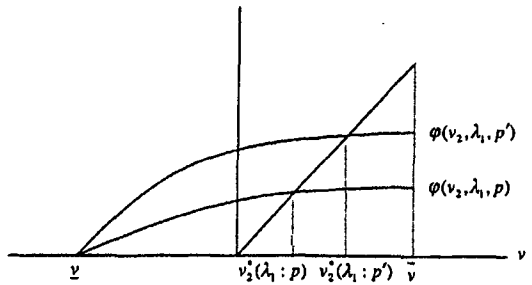


Figure 9

Effect of an Increase in p

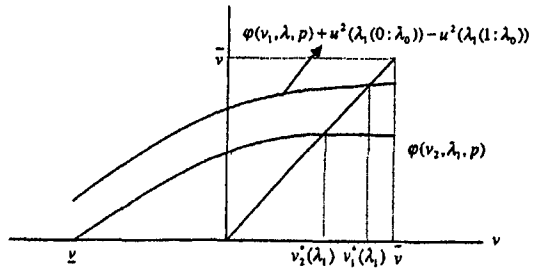


Lemma 4: $v_1^*(\lambda) > v_2^*(\lambda)$ for any $\lambda \in (0, 1)$.

Proof: Since $\lambda_1(1; \lambda_0) > \lambda_1(0; \lambda_0)$, we have $U^2(\lambda_1(0; \lambda_0)) > U^2(\lambda_1(1; \lambda_0))$ by lemma 3. This implies that $v_1^*(\lambda) > v_2^*(\lambda)$ from equation (9). (See figure 10.) *Q.E.D.*

Figure 10

Comparison of Incentives for Deviance in Two Periods



Finally, we have the following important lemma and proposition.

Lemma 5: There exists a $\bar{\lambda} > 0$ such that, for all $\lambda \in (0, \bar{\lambda})$, $\varphi(v, \lambda, p)$ is strictly increasing in λ for all $v \in V$.

Proof: See the appendix.

Proposition 7: $v_2^*(\lambda)$ is strictly increasing in λ, for all $\lambda \in (0, \bar{\lambda})$ for some $\bar{\lambda} > 0$.

Proof: We have $v_2^*(\lambda) = \varphi(v_2^*(\lambda), \lambda, p)$. Since $\varphi_2 > 0$ for all $\lambda \in (0, \bar{\lambda})$ for $\bar{\lambda}$ obtained from lemma 5, it is immediate that $\dfrac{\partial v_2^*}{\partial \lambda} > 0$ for such λ. *Q.E.D.*

This proposition suggests that, if λ is very small, it is possible that $v_2^*(\lambda_1(1, \lambda_0)) > v_2^*(\lambda_1(0, \lambda_0))$. That is, a rational individual who has an experience in engaging in deviant behavior may be less likely to deviate again, even if he is sophisticated in the sense that he takes into account the possibility of redeviation when he breaks the norm.

Before we close this section, we will discuss the probability that an individual may be completely stigmatized. It is transparent that an individual is completely stigmatized from period t on if and only if $v_t^* = \bar{v}$. This implies that a rational individual will not be completely stigmatized in equilibrium with positive probability. If $v_t^* = \bar{v}$ for some t, an act of deviant behavior would make the deviant completely stigmatized. However, the probability will be zero that an individual will deviate in equilibrium in that case. Therefore, we have the following proposition.

Proposition 8: It is only an individual known with certainty to be pathological that is completely stigmatized.

Tables 1 and 2 illustrate the equilibrium values for v_1^*, v_2^*, λ_1^*, and λ_2^* associated with various values of p under the assumption of uniform distribution.

Table 1

Equilibrium Values for $\lambda_0 = 0.2$, $w = 2$, and $\bar{v} = 5$

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\lambda_1(0; \lambda_0)$	0.1910	0.1807	0.1689	0.1553	0.1394	0.1207	0 .0985	0.0718	0.0395
$\lambda_1(1; \lambda_0)$	0.3479	0.3490	0.3503	0.3518	0.3535	0.3556	0.3581	0.3612	0.3650
$\lambda_2(0; \lambda_1(0))$	0.1823	0.1807	0.1420	0.1195	0.0956	0.0710	0.0467	0.0244	0.0072
$\lambda_2(1; \lambda_1(0))$	0.3346	0.3490	0.3029	0.2823	0.2573	0.2265	0.1881	0.1396	0.0780
$\lambda_2(0; \lambda_1(1))$	0.3348	0.3206	0.3039	0.2839	0.2597	0.2299	0.1924	0.1445	0.0822
$\lambda_2(1; \lambda_1(1))$	0.5373	0.5403	0.5439	0.5482	0.5534	0.5598	0.5677	0.5777	0.5907
$v_1^*(\lambda_0)$	0.3138	0.3366	0.3627	0.3929	0.4281	0.4697	0.5192	0.5788	0.6511
$v_2^*(\lambda_1(0))$	0.3045	0.3142	0.3218	0.3257	0.3234	0.3110	0.2828	0.2303	0.1414
$v_2^*(\lambda_1(1))$	0.4049	0.4394	0.4801	0.5286	0.5874	0.6598	0.7505	0.8663	1.0169

Table 2

Equilibrium Values for $\lambda_0 = 0.5$, $w = 5$, and $\bar{v} = 10$

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\lambda_1(0;\lambda_0)$	0.4851	0.4675	0.4462	0.4202	0.3877	0.3465	0.2929	0.2219	0.1268
$\lambda_1(1;\lambda_0)$	0.6903	0.6927	0.6956	0.6992	0.7037	0.7096	0.7174	0.7281	0.7430
$\lambda_2(0;\lambda_1(0))$	0.4703	0.4352	0.3936	0.3442	0.2863	0.2202	0.1486	0.0790	0.0236
$\lambda_2(1;\lambda_1(0))$	0.6777	0.6648	0.6488	0.6282	0.6004	0.5611	0.5015	0.4054	0.2463
$\lambda_2(0;\lambda_1(1))$	0.6777	0.6650	0.6493	0.6295	0.6036	0.5684	0.5176	0.4381	0.2971
$\lambda_2(1;\lambda_1(1))$	0.8282	0.8309	0.8343	0.8385	0.8437	0.8505	0.8597	0.8725	0.8916
$v_1^*(\lambda_0)$	1.0256	1.1258	1.2467	1.3949	1.5799	1.8157	2.1227	2.5309	3.0808
$v_2^*(\lambda_1(0))$	1.0370	1.1482	1.2761	1.4196	1.5704	1.7044	1.7641	1.6319	1.1135
$v_2^*(\lambda_1(1))$	0.7522	0.8298	0.9251	1.0451	1.2007	1.4108	1.7102	2.1718	2.9724

5 *Discussions*

5.1 Group Variations and Statistical Discrimination

Certain deviant behavior is not just individualistic. The rates of violent crimes in some age groups are higher than in other groups.[18] Specifically, male teenagers, who represent 3.3% of the U.S. population, commit 5.6% of its total homicides (Fox [1993]). Also, some kinds of criminal offenders are disproportionately drawn from certain racial and ethnic minorities. Wolfgang [1958] found that the homicide rate for blacks was four times as high as for whites in Philadelphia, and other researchers have reported similar findings for other cities (Reiss Jr. and Roth [1993]). These phenomena may provide a rationale for statistical discrimination. However, it is also obvious that such biases against those subgroups will most likely reinforce the incentive of the members within them to engage in deviant behavior. Our result (proposition 4) suggests that it may be better to adopt different deviance-deterring policies for those subgroups whose deviance rates are relatively high and relatively low, i.e., whose values of λ are different, whether for biological or for environmental reasons.

5.2 Ex Ante and Ex Post Perception in the Cost of Deviance

We have assumed that the distribution of v and the magnitude of the social sanction w are constant over time and that they are common knowledge. In general, however, they may change after the deviance. For example, an ex-criminal may have lower economic and psychological costs of implementing his criminal intent

[18] A subgroup of a society may develop its own norms that are not shared by the society of which it is a part. Such a subgroup is called a *subculture*. That is, a subculture is a culture within a culture. Thus, seemingly deviant behavior by the members of a subgroup may be the consequence of obeying its own norms.

because he has obtained knowledge about techniques of committing crimes from his past experience and, besides, he has become desensitized to anxiety, the fear of being apprehended, remorse, etc. This may shift the distribution of v downward once a crime is committed. Or it may be that an individual has an uncertain estimate of w. In particular, a potential deviant has a tendency to overestimate w before he first violates the norm. However, once he exhibits deviant behavior, he tends to begin underestimating it. This is because he does not want to think of himself as *persona non grata*. This may be one factor that makes an ex-deviant more likely to redeviate in many situations.

5.3 Observability and Heterogeneous Beliefs

In the real world, all of the community members cannot observe all the deviant actions. Accordingly, their perceptions of the potential deviant's mentality may not be the same, contrary to the assumption of this model that all of them observe the deviant's action and form one and the same posterior belief by Bayes's law.

In general, a more serious violation, usually involving a higher w, can be observed by more people. Bribery of an influential political figure hits the headlines, while a similar case wherein a policeman of no distinction is involved does not attract much attention.

Also, this may explain why a less sociable person is inclined to become a deviant. Since the behavior of such a person is rarely observable, his deviant behavior will not increase $\hat{\lambda}$ much. This, together with his small future payoffs from social interactions, will give him a strong incentive to deviate, which is the implication of proposition 3.[19]

6 *Stigma Effect and Policy Implications*

Stigmatization *ex ante* has the positive effect of deterring deviance, while *ex post* it has the negative effect of making the stigmatized person not hesitant to redeviate. That is, it can discourage current deviation, but it also encourages future deviation. For a larger positive effect to occur, a deviation must increase $\hat{\lambda} - \lambda$ greatly. Therefore, the tendency to hesitate to commit another deviation for fear of being stigmatized can appear only in the early stages of a deviant career.

Now, we may obtain some policy implications in light of the nature of the stigma effect. Suppose the objective of the government is to maximize the social welfare according to some appropriate definition. If the current crime rate is excessively high and a large proportion of crimes are committed by ex-convicts, one effective way to reduce the crime rate will be to alleviate the stigma effect. Meanwhile, stigmatization is possible only in the case that the public can observe

[19] Nonsociableness may be a consequence of deviant behavior as well as its cause. Persons enjoying deviant behavior, such as addiction or homosexuality, may avoid social gatherings and maintain an isolated status.

the action of a criminal. Thus, one way of mitigating the stigma effect might be to obliterate a criminal record after a certain period of nondeviation so as to keep the public from access to it. Of course, the adoption of this policy needs some caution, because it may put innocent people at higher risks of becoming victims of crimes. To be cautious in recommending this policy, let us consider a more elaborate model that explicitly deals with externalities generated by deviant behavior. Suppose, given λ, a potential deviant's neighbors first take precautions against possible losses due to his deviance. Let c be the precaution level. For simplicity, assume that the short-run gain of the deviant, v, is independent of c and that precautions reduce losses. That is, letting $L(c)$ be the losses when a precaution level c is taken, we assume that $L'(c) < 0$ and that $L''(c) > 0$. After precautions are taken, the potential deviant decides whether or not to follow the norm, and as a consequence, $\hat{\lambda}$ is updated.

The potential deviant's behavior is determined independently of the precaution level, due to the assumption that c does not affect v. Neighbors choose c to minimize their expected total cost $\varrho L(c) + c$, where $\varrho = \lambda + (1 - \lambda)(1 - F(v^*))$ is the probability that the individual perpetrates a deviant act. Notice that the solution for the optimization problem, c^*, is increasing in ϱ if it is an interior solution. If the government finds an ex-convict to whom the stigma is attached so that he has no other choice but to commit another crime for a living, the policy of obliterating his criminal record will reduce the likelihood of recidivism. Knowing this, neighbors will reduce the precaution level. Therefore, in this case, this policy recommendation is clearly welfare-improving. The problem with this policy, however, lies in the case of a pathological criminal. In that case, the policy will not help deter the criminal acts and will only reduce the precaution level people take. As a consequence, larger social losses are expected. Therefore, to ensure the effectiveness of this policy, it is essential to tell to what extent the potential criminal is habitual. Mandating a certain period of no criminal activity before expunging the record may help screen out habitual criminals.

Rasmusen seems somewhat skeptical of the policy of keeping criminal records secret. He argues that the policy may have the effects of preventing discrimination against ex-criminals as workers and of raising their wage (and thus the opportunity cost of crime), thereby lowering the incentive to commit a crime. But he also argues that the latter effect can be achieved by a direct policy of keeping records open, but subsidizing the wage of ex-criminals, which would match them better with jobs. It may be true that this alternative policy would not distort the labor market, but it would not discourage secondary deviance very much, while encouraging primary deviance.

Analogously, the Korean government recently decided to expunge a citizen's divorce record from his or her family register. Since a divorce is still regarded as shameful in Korea, this reform has been welcomed by many suffering divorced persons (particularly women) who could no longer be stigmatized. However, it may also have the effect of increasing the divorce rate. If the Korean government does worry about the currently growing divorce rate and consider it as a social concern, this policy measure may be regarded as pernicious.

7 Unofficial Stigmatization versus Official Punishment

The incentive to break social norms can be controlled either by informal sanctions of stigmatization or by formal sanctions of legal punishment.

There is some literature on deterring criminal behavior of repeat offenders by legal sanctions. RUBINSTEIN [1979] is the first paper dealing with the optimal sanctions for repeat offenders. He showed that it is sometimes optimal not to punish first-time offenders. Subsequent researches, including POLINSKY AND RUBINFELD [1991] and CHU, HU, AND HUANG [2000], addressed the same issue on different assumptions. In particular, the latter workers showed that it is always optimal to punish repeat offenders more severely than first-time offenders, as long as one considers the possibility of the erroneous conviction of innocent offenders.

Those papers resemble this one in discussing the intertemporal sanctioning pattern for a repeat deviant (or criminal). However, the official punishment pattern for a repeat criminal over time is determined *ex ante* by the legal authority's effort to minimize the social cost associated with committing a crime, while community members' informal sanctioning for deviant behavior is determined *ex post* according to a natural belief formation procedure – the so-called Bayesian updating rule – as a result of an individual's deviant behavior.

Deviant behavior is in general not illegal *per se*, so that it cannot be subject to official punishment, but it is possible to discipline a potential deviant through informal sanctioning – that is, through stigmatization by the community members.

8 Conclusion

Contrary to the existing literature on the labeling theory of deviance that explicitly assumes a nonsmooth, one-time stigmatization of a deviant, this paper explores by what process a deviant is stigmatized and stresses that stigmatization possibly has a positive effect of deterring deviance as well as its well-known negative effect. More importantly, our paper provides a theoretical analysis of the relationship between the experience of having deviated and the incentive for deviation, and suggests the possibility that an ex-deviant may be less likely to deviate, as opposed to the conventional wisdom that the reverse will be the case.

Although some descriptive illustrations have been provided, this paper does not contain an empirical analysis supporting the theoretical finding, in view of the difficulty in collecting a large enough sample of data for meaningful empirical analysis. But we are sure that empirical research testing the hypothesis put forward in this paper will have substantial academic value and have a significant effect on policy making.

This model can be extended in various ways. First, we may consider a model in which the chance to make a decision is stochastically given at each period $t = 1, 2, \ldots$ Then, a deviant individual is sure to be completely stigmatized at some time, although that will not occur in equilibrium with positive probability. However, no

other results we have obtained will be qualitatively affected. Alternatively, we can consider a more sophisticated model in which a sequence of chances of getting short-run gains from violating the norm are given according to a counting process (presumably a Poisson process) and the public keeps updating its posterior belief every time an individual violates the norm. Also, it will be worthwhile pursuing richer models incorporating various features not captured in this model, as discussed in section 5.

Appendix

A.1 Proof of Proposition 4

Given v, let $1 - F(v) \equiv q$ and define $\Lambda = \hat{\lambda} - \lambda = \frac{\lambda}{\lambda + (1-\lambda)q} - \lambda$. If $q = 1$, i.e., $v = \underline{v}$, then Λ is identically zero, so that $\frac{\partial \Lambda}{\partial \lambda} = 0$ for any λ. If $q \neq 1$, then $\frac{\partial \Lambda}{\partial \lambda} > 0$ is equivalent to $\psi(\lambda; q) < 0$, where $\psi(\lambda; q) = (1-q)\lambda^2 + 2q\lambda - q$.

First, since $\psi(0; q) = -q < 0$, $\exists \lambda_\varepsilon(v)\ (> 0) \ni \forall \lambda < \lambda_\varepsilon(v)$, $\psi(\lambda; q) < 0$ for any given v, and thus we have $\frac{\partial \Lambda}{\partial \lambda} > 0$ for any $\lambda < \lambda_\varepsilon(v)$. Thus, $\phi(v)$ schedule is shifted upwards with an increase in λ, so that v^* is increased as λ is increased.

On the other hand, if $\psi(\lambda; 1) = 2\lambda - 1 > 0$, i.e., $\lambda > \frac{1}{2}$, we have $\psi(\lambda; q) > 0$ for any $q \in [0, 1]$, since $\psi(\lambda; 0) = \lambda^2 > 0$ and $\frac{\partial \psi}{\partial q} = -(\lambda - 1)^2 < 0$, and thus $\frac{\partial \Lambda}{\partial \lambda} < 0$ for any $q \in [0, 1)$. Therefore, for any $v \neq \underline{v}$ we have $\frac{\partial \Lambda}{\partial \lambda} < 0$, which implies that an increase in λ shifts $\phi(v)$ schedule downward, so that v^* is decreased with an increase in λ. *Q.E.D.*

A.2 Proof of Lemma 5

A bit of calculus gives $\varphi(v, \lambda, p) = h(\lambda, v, p) - g(\lambda, v, p) = w \frac{F(v)}{1-p} \frac{\lambda(1-\lambda)}{\Delta_1 \Delta_2}$, where $\Delta_1 = \lambda + (1-\lambda)(1 - F(v))$ and $\Delta_2 = 1 + \frac{1}{1-p}(1-\lambda)F(v)$. Then, $\varphi_2 = w \frac{F(v)}{1-p} \frac{\Xi}{\Delta_1^2 \Delta_2^2}$, where $\Xi = \{(1+\lambda)(1-2\lambda) - \lambda(2-3\lambda)F(v)\}\Delta_2 + \lambda(1-\lambda)\Delta_1 \frac{F(v)}{1-p}$. Therefore, $\exists \lambda_\varepsilon > 0$ such that, $\forall \lambda \in (0, \lambda_\varepsilon)$, $\Xi(\lambda) > 0$, so that $\varphi_2 > 0$. *Q.E.D.*

References

AKERLOF, G. [1980], "A Theory of Social Custom, of Which Unemployment May be One Consquence," *Quarterly Journal of Economics*, 94, 749–775.

-- [1997], "Social Distance and Social Decisions," *Econometrica*, 65, 1005–1027.

BECKER, G. [1968], "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76, 169–217.

BECKER, H. [1963], *The Outsiders*, University of Chicago Press: Chicago, IL.

BENSING, R., AND O. SCHROEDER [1960], *Homicide in an Urban Community*, Charles C. Thomas: Springfield, IL.

BERNHEIM, D. [1994], "A Theory of Conformity," *Journal of Political Economy*, 102, 841–877.

BRAITHWAITE, J. [1989], *Crime, Shame, and Reintegration*, Cambridge University Press: Cambridge.

CHERLIN, A. [1978], "Remarriage as an Incomplete Institution," *American Journal of Sociology*, 84, 634–650.

CHU, C., S. HU, AND T. HUANG [2000], "Punishing Repeat Offenders More Seriously," *International Review of Law and Economics*, 20, 127–140.

CLOWARD, R., AND L. OHLIN [1960], *Delinquency and Opportunity*, Free Press: New York.

COLEMAN, J. [1987], "Norms as Social Capital," pp. 133–155 in: G. Radnitzky and P. Bernholz (eds.), *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics*, Paragon: New York.

COOK, P., AND G. ZARKIN [1984], "Crime and the Business Cycle," Working Paper #84-02, Duke University, Durham, NC.

FOX, J. [1993], *Teenage Males are Committing Murder at an Increasing Rate*, National Center for Juvenile Justice: Pittsburgh, PA.

GLAESER, E., B. SACERDOTE, AND J. SCHEINKMAN [1996], "Crime and Social Interactions," *Quarterly Journal of Economics*, 111, 507–548.

MERTON, R. [1938], "Social Structure and Anomie," *American Sociological Review*, 3, 672–682.

POLINSKY, M., AND D. RUBINFELD [1991], "A Model of Optimal Fines for Repeat Offenders," *Journal of Public Economics*, 46, 291–306.

RASMUSEN, E. [1996], "Stigma and Self-Fulfilling Expectations of Criminality," *Journal of Law and Economics*, 39, 519–543.

REISS JR., A. J., AND J. A. ROTH [1993], *Understanding and Preventing Violence*, National Academy Press: Washington, DC.

RUBINSTEIN, A. [1979], "An Optimal Conviction Policy for Offenses that may have been Committed by Accident," pp. 406–413 in: S. Brams, A. Schotter, and G. Schwodiauer (eds.), *Applied Game Theory*, Physica-Verlag: Wurzburg.

SUTHERLAND, E. [1947], *Principles of Criminology*, 4th ed., Lippincott: Philadelphia, PA.

WITTE, A. D., AND H. TAUCHEN [1994], "Work and Crime: An Exploration Using Panel Data," NBER Working Paper #4794, NBER, Cambridge, MA.

WOLFGANG, M. [1958], *Patterns of Criminal Homicide*, University of Pennsylvania Press: Philadelphia.

Jeong-Yoo Kim
Department of Economics
Dongguk University
3-26, Pildong, Chungku
Seoul 100-715
Korea
E-mail: jyookim@dongguk.edu

Gang Lee
Department of Sociology and Anthropology
University of Texas at El Paso
Old Main, Room 317
El Paso, TX 79968-0558
USA
E-mail: lgang@utep.edu