Fine enough or don't fine at all

Chung-cheng Lin [a,b], C.C. Yang [a,b,*]

[a] *Institute of Economics, Academia Sinica, Nankang, Taipei 115, Taiwan*
[b] *Department of Public Finance, National Chengchi University, Wenshan, Taipei 116, Taiwan.*

Abstract

This paper explores the results discussed in "A fine is a price" by Gneezy and Rustichini [2000a. A fine is a price. Journal of Legal Studies 29, 1–17] regarding the effect of fines on parents who collect their children late from day-care centers. We suggest a complementary explanation, analyzing a formal model that qualifies but does not lose the predictive power of the deterrence hypothesis. Our main policy implication can be summarized as "Fine enough or don't fine at all." We suggest that the Gneezy and Rustichini finding may represent a more general result and that our formal model is potentially applicable to a variety of situations.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

The higher the price of a good, all else being equal, the lower will be the quantity consumed of that good. This negative relationship between price and quantity is known as the "law of demand" in economics. This "law" is asserted to be equally applicable to crime and punishment according to the deterrence hypothesis: the higher the cost of committing crimes, all else being equal, the lower will be the crime rate.[1]

---

* Corresponding author. Tel.: +886 2 27822791; fax: +886 2 27853946.
E-mail address: ccyang@econ.sinica.edu.tw (C.C. Yang).

[1] For an introduction to the deterrence hypothesis, see Cooter and Ulen (2000).

In a provocative article, Gneezy and Rustichini (2000a, hereafter G&R) challenged the predictive strength of the deterrence hypothesis. They conducted a field study involving a group of day-care centers. Some parents would arrive late to collect their children, forcing a teacher to stay after closing time and wait with the children concerned. G&R obtained two results that strongly contradicted the prediction of the deterrence hypothesis. First, the introduction of a relatively small but not insignificant monetary fine for parents who came late significantly increased the number of parents who came late.[2] Second, this number of late-comers remained higher and did not return to the pre-fine level after the fine was removed.[3]

G&R offered two models to explain their finding. Both models are built on the assumption that the environment in which agents play is defined by an incomplete contract; the introduction of a fine into this incomplete contract alters the information that the agents have, and hence changes the agents' perception of the environment. G&R argued that such a change in the perception of the environment may give rise to effects different from what might be expected from the "all else equal" clause included in the deterrence hypothesis.

G&R's first model is an application of game theory with asymmetric information. Through infinite plays, parents learn whether the owner of a day-care center is a severe type or a mild type. The introduction of an explicit fine completely reveals that the owner is a mild type; therefore, the fine is the worst that can be imposed on parents who come late. This information revelation induces an increase in the number of late-coming parents. After the fine is removed, parents still have the same information and hence have no reason to anticipate anything worse than the fine. As a result, the number of late-coming parents remains higher.

The second model is informal and is based on the concept of social norms. There are three norms involved. Before the introduction of a fine, the stay and wait action of teachers may be interpreted as a generous form of nonmarket behavior, and consequently, parents should be restrained from taking advantage of it (the first norm). Imposing the fine puts a price on the service and parents can buy the service from day-care centers after closing time as much as they find it convenient (the second norm). After the fine is removed, the service is simply perceived to be offered at zero price since there is a third norm: "Once a commodity, always a commodity."

There is no doubt that G&R's explanations are interesting and consistent with the finding from their field study. However, their explanations are not without weaknesses. In the case of the game theory model, one cannot rule out the possibility that the information concerning the type of the owner of a day-care center may be obtained by checking the operating history of that center. Indeed, if there exists a severe type of owner, it seems not unreasonable to speculate that many parents will engage in such checking before choosing a specific

---

[2] For the precise meaning of "a small but not insignificant monetary fine," see G&R.

[3] In the interpretation of their second result, G&R pointed out that (i) the number of late-coming parents remained higher than the pre-fine level, and (ii) this number remained stable after the fine was removed. In this paper, however, we shall emphasize point (i) only. In view of the fact that the observation period after the fine was removed is relatively short, our emphasis, we believe, is not incompatible with the G&R finding.

day-care center for their children.[4] As to the social norm model, involving three social norms to explain the finding may appear awkward. At the end of their paper, G&R concluded that they could not refute either of the two models on the basis of the data and the theory available. In spite of this conclusion, G&R (p. 15) also concluded: "A satisfactory answer will be interesting, however, because these facts are probably more than a curious finding."

In this paper we propose a complementary explanation in an attempt to gain further insights into the G&R finding. Our explanation is built on a formal model in which the concept of social norms and the psychological cost arising from the violation of social norms are linked together. Unlike the G&R social norm model, there is a unique social norm throughout our model that prescribes that individuals should be on time and not late. The introduction of a monetary fine for parents who come late is deemed to reduce the psychological cost arising from the violation of the social norm, and this, in turn, erodes the bite or effectiveness of the social norm against delinquency. The resulting erosion of the social norm provides us with a way of explaining why the introduction of a monetary fine may significantly increase the number of parents who come late. Without appealing to the norm of "Once a commodity, always a commodity," we are also able to explain within our coherent model why the number of late-coming parents remains higher and does not return to the pre-fine level after the fine is removed. Unlike G&R's own explanation, the predictive power of the deterrence hypothesis based on our model is qualified but not lost. The main policy implication from our analysis can be summarized by a simple sentence: "Fine enough or don't fine at all."[5] This "stick" policy prescription is symmetric to the "carrot" policy prescription reached by Gneezy and Rustichini (2000b): "Pay enough or don't pay at all." In conclusion, we suggest that the G&R finding is not a singleton but may represent a more general result. We also suggest that the formal model we develop in this paper is potentially applicable to a variety of situations.

The remainder of this paper consists of five parts. Section 2 introduces our model. Section 3 analyses the model, paving the way for our explanation of the G&R finding in Section 4. Section 5 provides some discussion and Section 6 concludes.

2. Model

Consider a community in which parents send their children to day-care centers. Parents can choose to collect their children either before the day-care centers close (on time) or

---

[4] G&R (p. 4) mentioned: "All of these centers are located in the same part of town, and there is no important difference among them." Since day-care centers are privately owned, one may wonder whether the severe-type owner will ever implement severe actions such as the exclusion ("Kicking out") of the child from the day-care center. The exclusion means that customers are expelled. However, the cost that will be imposed on the parents in question seems mild since these parents can transfer their children to another nearby, not so different day-care center.

[5] After we finished writing the paper, we found that a similar explanation for the G&R finding was independently proposed in a rather broad context by Fehr and Falk (2002). Although similar, some important differences remain. First, we provide a formal model with a micro foundation, whereas the Fehr and Falk paper did not. Second, and perhaps more importantly, we demonstrate that the predictive power of the deterrence hypothesis is qualified but not lost, while the Fehr and Falk paper seems to concur with the G&R view.

after (behind time or coming late). Let the choice be denoted by t, which for simplicity is assumed to take on the value 0 if on time and 1 if behind time. Suppose that parents will have an extra unit of time to use at their own convenience if they choose to come late. Then our setup implies that this extra unit of time will not be available to parents who choose to be on time.

Parents maximize a utility function:

$$U = m + vt - C \tag{1}$$

where m is income, v is the value of time, and C is the psychological cost associated with the choice of t: $C = 0$ if $t = 0$ but $C > 0$ if $t = 1$. The C term in (1) plays a key role in this paper. Our modeling strategy is to link it with social norms. In the remainder of this section, we discuss the modeling of the C term in detail.

In a seminal paper, Akerlof (1980) put forth the idea that there exist a variety of codes of behavior (social customs or social norms) in the community. He argued that these social customs or social norms might persist without erosion even if obeying them is disadvantageous to people's pecuniary payoffs. Romer (1984), Booth (1985), Naylor (1989), Lindbeck et al. (1999) and others have extended the argument further. Following this segment of the literature, it is assumed that there exists a social custom or social norm (a code of behavior or honor) in the community in question. This norm, shared commonly among people in the community, prescribes that individuals should be on time and not behind time. This norm applies of course to the case where parents should collect their children on time at day-care centers. Casual evidence and daily experience indicate that many societies do have this "on time" norm.[6] The focus of this paper is not on why such a norm is established in the first place, but rather on its implications.

Norms such as individuals needing to be on time and not behind time are hardly suitable for embodiment in law. Without the enforcement of the state, it seems clear that there must exist informal social sanctions to enforce these norms to ensure their survival. According to Posner and Rasmusen (1999), there are various sanctions that enforce norms. For simplicity and ease of exposition, we focus on the sanctions of guilt and shame and summarize the resulting guilt and shame imposed upon a violator of the "on time" norm by his or her psychological cost incurred. This psychological cost is represented by the C term in (1).

Guilt is an internal sanction. It may occur even if the violator will never be caught for sure. By contrast, shame is basically an external sanction. It occurs because other people find out about the violation and because the violator feels lowered in his or her own eyes or in the eyes of other people. Although guilt is an internal sanction, it is incurred as a result of the process of socialization such as through education and upbringing. As an external sanction, shame must have to do with the actions or reactions of other people, say, group disapproval, peer pressure, or in the extreme, ostracism.[7] Since parents who come late are always known or detected, both guilt and shame are likely to result in our context. To be

---

[6] "Be punctual as to time" is listed as an important form of etiquette and one of the commonplace day-to-day obligations in Bennett (1993, p. 201). The popularity of the dictum "time is money" may also reflect to some extent the wide acceptance of the "on time" norm around the world.

[7] For details on guilt, shame and other sanctions, see Posner and Rasmusen.

specific, the psychological cost C associated with guilt and shame in (1) is specified as

$$C = \lambda c(f, x); \quad c_f = \frac{\partial c}{\partial f} < 0, \quad c_x = \frac{\partial c}{\partial x} < 0 \qquad (2)$$

where λ is an idiosyncratic sensitivity indicator that may capture the degree of personal belief in the "on time" code,[8] and $c(\cdot)$ denotes a common psychological cost function that is decreasing with respect to both the fine f and the portion of late-coming parents x.

Parents in our model are heterogeneous in the sense that λ is idiosyncratic and varies across parents. For simplicity, it is assumed that the distribution of λ is uniform with support on [0,1]. Given $c(\cdot)$, the higher the value of λ, the higher will be the psychological cost incurred. It should be noted that the psychological cost inflicted on parents who come late in (2) depends not only on their own *subjective* view of the code of behavior (i.e. λ) but also negatively on the *objective* portion of parents who violate the norm (i.e. x). For the "on time" norm to be social, people in the community must share and sustain it. We assume the dependence of c on x in (2). As to the setup $c_x < 0$, it is in line with the emphasis in the social norm literature that the bite or effectiveness of social norms against their violators will become less intense if the violation becomes more prevalent.[9]

G&R put forth the novel idea of "a fine is a price" as a monetary fine is introduced and imposed on parents who come late. "A fine is a price" is viewed as a social norm in G&R's (p. 14) social norm model: "The two labels of the payment as price or fine are equivalent." Once this norm is formed, parents feel justified to buy the service from day-care centers after

---

[8] Following the seminal work of Akerlof, social custom models often make a dichotomy between those who believe in the code of behavior and those who do not. The key distinction between believers and non-believers in this Akerlovian type of model is that believers suffer from internal sanctions (such as guilt) merely by disobeying the code, while non-believers do not. The psychological cost term C in our model includes both guilt (an internal sanction) and shame (an external sanction). Through the single sensitivity indicator λ, we do not differentiate between the idiosyncratic sensitivity toward guilt and that toward shame. As a result, the personal belief in the code of behavior in our model is no longer dichotomous but instead becomes continuous. Note that the distribution of the personal belief is assumed fixed and does not evolve in our model. Since G&R conducted their experiment over a period of only 20 weeks, this assumption does not seem inappropriate as far as explaining the G&R finding is concerned.

[9] See Akerlof and his followers. It should be pointed out that the focus of this segment of the literature is on the so-called "global externalities": externalities work through the effect of the global or economy-wide average on individual behavior. We follow this literature, and the portion of late-coming parents x, incorporated via the C term in (2), represents the global or economy-wide average in our model. There is a different segment of the literature where the focus is on the so-called "local externalities": externalities work through the effect of individual agents on each other's behavior. Modeling strategy along local externalities is quite distinct from that along global externalities. For a local model and a discussion on the distinction between local and global externalities, see Glaeser et al. (1996). In the G&R experiment, the size of day-care centers (in terms of the number of children) ranges from 28 to 37, and the average size is around 34. These numbers may not be large, but they are not small either. Our global approach should not be a bad first-order approximation if parents are not well enough acquainted with each other and interact somewhat anonymously. Alternatively, since coming late involves low benefits and low costs and since the decision itself may be costly, parents who decide to be on time or behind time may fall back on simple rules of thumb as captured by our global approach. For a defense of model-building strategy that postulates behavior based on simple rules of thumb, see Akerlof and Yellen (1991).

closing time to the extent that they find it convenient. Here, the introduction of a monetary fine simply opens up the possibility of a quid pro quo: parents pay a "price" for the service provided by day-care centers after closing time. Although a "price" is paid for the service, price and fine may not be equivalent in our model since the price includes the guilt and shame inflicted as well as the fine imposed.

In the field of psychology, there is the so-called "equity theory."[10] It basically says that in interpersonal or social exchange, the perceived value of "inputs" will tend to equal the perceived value of "outcomes" in the subjective sense. This theory implies, for example, that workers will reduce their labor effort to achieve an effort–remuneration balance if they think subjectively that they are underpaid. Empirical studies are on balance strongly supportive of this theory according to Akerlof and Yellen (1990). On the basis of this theory, it is arguable that, when the fine is introduced, the guilt and shame associated with coming late will be lessened since the "outcome" remains unchanged (late-coming parents still "enjoy" the same value of time v), but there comes an additional "input" of the fine. Just as workers who reduce their labor effort to achieve an effort–remuneration balance, late-coming parents "adjust" their guilt and shame *downward* to achieve a balance between the perceived value of the total "input" (fine plus guilt and shame) and the perceived value of the "outcome" (the value of time v). As the fine imposed gets higher, it is likely, according to the equity theory, that the extent of the guilt and shame associated with coming late will decrease. Put simply, the introduction of a monetary fine may be thought of as "lessening" a late-coming parent's guilt and shame arising from the violation of the "on time" norm. The extent of the "lessening" will become larger as the fine imposed is raised. This leads to the setup $c_f < 0$ in (2).[11]

This completes the description of our model.

3. Analysis

According to our setup in the previous section, the utility of a type λ parent equals

$$U^0 = m \quad \text{if } t = 0 \text{ (on time)}; \tag{3a}$$

$$U^1 = m + v - f - \lambda c(f, x) \quad \text{if } t = 1 \text{ (behind time)} \tag{3b}$$

where $f = 0$ in (3b) if there is no fine for late-coming parents.

A type λ parent will choose to come late if $U^1 \geq U^0$, that is, if the following inequality holds:

$$v - f - \lambda c(f, x) \geq 0. \tag{4}$$

---

[10] See Adams (1963).

[11] Following the equity theory, the guilt and shame associated with coming late will be lessened but will not disappear as long as the monetary fine imposed f is smaller than the value of time v. It is worth mentioning that a version similar to the equity theory appears as the "social exchange theory" in the field of sociology; see Akerlof and Yellen (1990).

From (4), the marginal type of parents ($\hat{\lambda}$) who are merely indifferent between being on time and behind time is given by

$$\hat{\lambda} = \frac{v - f}{c(f, x)} \tag{5}$$

with

$$\hat{\lambda}_f = \frac{\partial \hat{\lambda}}{\partial f} = -\frac{1}{c} - \frac{(v - f)c_f}{c^2} \gtreqqless 0; \tag{5a}$$

$$\hat{\lambda}_x = \frac{\partial \hat{\lambda}}{\partial x} = -\frac{(v - f)c_x}{c^2} > 0; \tag{5b}$$

$$\hat{\lambda}_{xx} = \frac{\partial^2 \hat{\lambda}}{\partial x^2} = -\frac{(v - f)c_{xx}}{c^2} + \frac{2(v - f)(c_x)^2}{c^3} \gtreqqless 0. \tag{5c}$$

Parents with $\lambda \le \hat{\lambda}$ will choose to come late, while those with $\lambda > \hat{\lambda}$ will choose to be on time. In (5), the value of time v is assumed to be larger than the fine f; otherwise, no parents will choose to come late according to (5) (i.e. $\hat{\lambda} \le 0$). The same assumption is also imposed in G&R.

The sign of $\hat{\lambda}_f$ may be positive or negative, depending on the two counterbalancing effects captured by the two terms in (5a). The first term is the standard deterrence effect. It is negative because a larger fine raises the monetary cost of coming late. The second term is the social norm-psychological cost effect. It is positive because a larger fine will reduce the psychological cost associated with the guilt and shame arising from coming late. The number of parents who choose to come late may thus increase or decrease as a result of imposing fines according to (5a).

More precisely, we have from (5a) that $\hat{\lambda}_f > 0$ if and only if

$$\varepsilon > \frac{f}{v - f} \tag{6}$$

where $\varepsilon = -(\partial c / \partial f)(f/c)$ denotes the elasticity of the reduction in the psychological cost through the imposition of fines. Given ε and v, we see from (6) that there exists a fine level, $f_0 = \varepsilon v/(1 + \varepsilon)$, such that the inequality (6) holds if and only if $f < f_0$. This in turn implies that the social norm-psychological cost effect will dominate the deterrence effect (i.e. $\hat{\lambda}_f > 0$ in (5a)) if and only if $f < f_0$. It is important to note that $f_0 < v$.[12]

---

[12] Notice that the elasticity, ε, may depend on the fine imposed, f, and the portion of late-coming parents, x. For simplicity, however, the elasticity ε will be assumed to be a constant in our argument. This constant-elasticity assumption in the argument is not uncommon in economics. Moreover, from the viewpoint of econometrics, the elasticity ε will be estimated to be some constant as a rule. Let $c(f, x) = g(x)/f$ with $g' < 0$. This is a plausible functional form for the common psychological cost term $c(f, x)$. It can be checked in this case that ε is a constant equal to 1 and $f_0 = v/2$. For those who do not like the idea that the elasticity ε is assumed to be a constant, it might be useful to proceed directly with this case in the argument.

The outcome $\hat{\lambda}_x > 0$ in (5b) captures the "snowballing" effect: the higher the portion of late-coming parents at the status quo, the larger the number of parents who will choose to come late rather than be on time. When the behavior of coming late is more prevalent, the "on time" norm will become less effective against late-coming parents, and consequently, the more prevalent that coming late is, the greater the extent that coming late will be intensified.

Since c_{xx} could be positive as well as negative, the sign of $\hat{\lambda}_{xx}$ in (5c) is ambiguous a priori. To pin down the reasonable possibilities, one may think of (5) as a production function where $\hat{\lambda}$ is an "output" and x is an "input." There are two typical relationships between an output and an input in economics. First, the marginal "contribution" of the input to the output is increasing at the initial stage, but it will eventually begin to decrease once the "law of diminishing returns" sets in. Second, the "law of diminishing returns" sets in immediately so that the marginal "contribution" of the input to the output will be decreasing from the beginning.[13] We will consider both relationships for $\hat{\lambda}$ and x.

Since λ is assumed uniformly distributed with support on [0,1], by the definition of $\hat{\lambda}$, we also have

$$\hat{\lambda} = x. \tag{7}$$

Given a portion of late-coming parents, there is a corresponding portion of parents who will choose to come late. However, the resulting $\hat{\lambda}$ may not be consistent with the given x. Eq. (7) simply imposes the consistency condition. It is clear that, except for corner solutions, an equilibrium x must satisfy (5) and (7) simultaneously.

We are now ready to explain the two main results of G&R's field study: (i) the introduction of a monetary fine for parents who came late significantly increased the number of parents who came late, and (ii) this number of late-comers remained higher and did not return to the pre-fine level after the fine was removed.

4. Explanation

There are three periods in the G&R field study: the initial period where no fine is imposed, the second period where a relatively small but not insignificant fine is imposed, and the third period where the fine imposed is removed. Our explanation corresponds to these three periods and is illustrated with several figures.

4.1. No fine for late-coming parents

In Fig. 1, the locus *XX* ($f = 0$) stands for the functional relationship between $\hat{\lambda}$ and x as expressed in (5) when no fine is imposed for late-coming parents. The slope of *XX* is positive because of (5b). Note that the slope of *XX* is increasing when x is low but begins decreasing when x is high. The shape of this graph is due to our assumption that the marginal

---

[13] For a clear exposition of the law of diminishing returns, see Stigler (1987). For a forceful argument for the applicability of the law to every conceivable subject, see Mundell (1968).
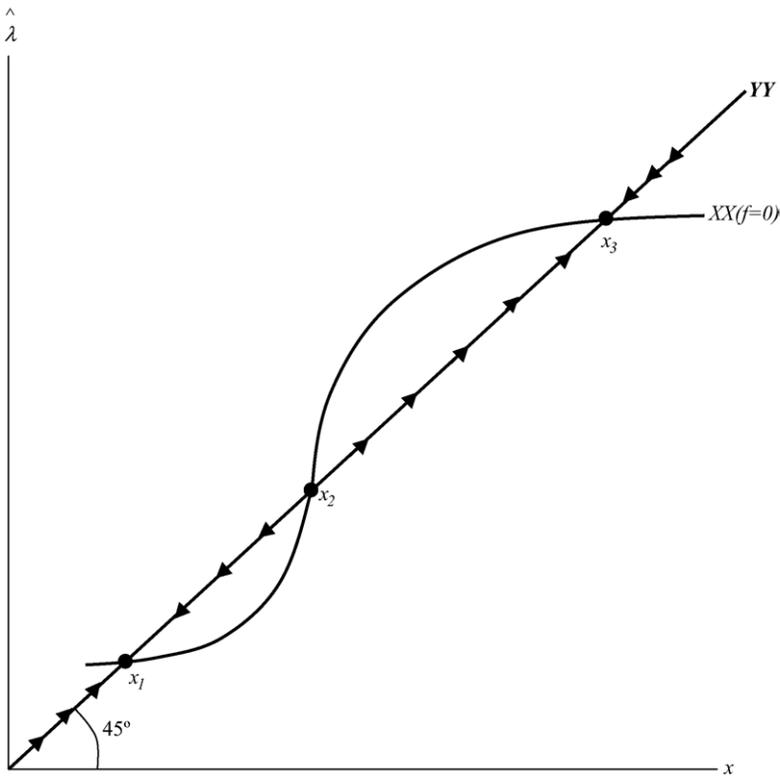
Fig. 1. No fine.

"contribution" of x to $\hat{\lambda}$ is increasing at the initial stage, but the marginal "contribution" will eventually begin to decrease once the "law of diminishing returns" sets in; that is, $\hat{\lambda}_{xx} > 0$ when x is low, but $\hat{\lambda}_{xx} < 0$ when x is high.[14]

In Fig. 1, the locus YY traces the relationship between $\hat{\lambda}$ and x as expressed in (7). It is obvious that the slope of YY equals 1.

Given any x at the status quo, the portion of late-coming parents will be increasing if $\hat{\lambda} > x$, but it will be decreasing if $\hat{\lambda} < x$. The reason behind this result is intuitive. When $\hat{\lambda} > x$, the portion of parents who will choose to come late is higher than the status quo portion of parents who come late. As a result, the actual portion of late-coming parents will be increasing. When $\hat{\lambda} < x$, the opposite occurs. The arrows in Fig. 1 summarize the movement of x.[15]

---

[14] This result is compatible with the assumption that $c_{xx} < 0$ when x is low but $c_{xx} > 0$ when x is high; that is, the law of diminishing returns for $c(\cdot)$ sets in after a while, not from the beginning. Since the sign of c_x is negative but that of $\hat{\lambda}_x$ is positive, the functional shape of c is consistent with that of $\hat{\lambda}$ under this assumption. We will discuss the other case where the law of diminishing returns sets in from the beginning in Section 5.

[15] A simple dynamics that captures the movement shown in Fig. 1 is $\dot{x} = \alpha(\hat{\lambda} - x)$ where α is a positive scalar.

There are three equilibria that satisfy (5) and (7) simultaneously in Fig. 1, that is, points x_1, x_2 and x_3. However, as the arrows indicate, only x_1 and x_3 are stable equilibria.[16] If the status quo portion of late-coming parents is not large enough before the fine is imposed, then the equilibrium will be at point x_1 initially. The data collected in the G&R study indicate that the portion of late-coming parents is relatively low before the fine is imposed.

4.2. A relatively small but not insignificant fine is imposed

Let us introduce a relatively small but not insignificant monetary fine and impose it on late-coming parents. Given ε and v, we know from (6) that there exists a fine level, f_0, such that the inequality (6) holds and $\hat{\lambda}_f > 0$ in (5a) if and only if $f < f_0$. Since the fine imposed is relatively small, it is arguable that the inequality (6) holds and $\hat{\lambda}_f > 0$ in (5a) in our context. This indicates that the locus XX associated with the relatively small fine will be located higher than the locus XX ($f = 0$) shown in Fig. 1.

Although the fine imposed is relatively small, it is not insignificant. This implies that the upward shift of the locus XX may not be insignificant either. Consider the locus XX ($f_1 > 0$) in Fig. 2. The unique feature of this locus is that it is tangential to the locus YY at the point y_1. Suppose that the fine imposed is no less than the fine associated with the locus XX ($f_1 > 0$) but that it is not large enough so that the social norm-psychological cost effect still dominates the deterrence effect (i.e. $f_1 \le f < f_0$ so that $\hat{\lambda}_f > 0$). Then the equilibrium will jump from x_1 to x_4 or even higher after the imposition of the fine (see Fig. 2). In other words, we have obtained a result similar to that in the G&R study: the introduction of a monetary fine for late-coming parents increases significantly the number of late-coming parents.

Note that, except for $\hat{\lambda} = x$ at point y_1, $\hat{\lambda} > x$ for all x satisfying $x < x_4$ for the locus XX ($f_1 > 0$) in Fig. 2. That is, after the introduction of a monetary fine, the portion of parents who will choose to come late is almost always larger than the status quo portion of late-coming parents if the latter is smaller than x_4. This triggers a bandwagon effect to increase the actual portion of late-coming parents, and the increase will not stop until the actual portion has reached x_4. The fine, f_1, associated with the locus XX ($f_1 > 0$) in Fig. 2, represents a threshold level in our model. As long as the social norm-psychological effect dominates the deterrence effect so that $\hat{\lambda}_f > 0$, the introduction of a monetary fine that is no less than the threshold level f_1 will generate a substantial increase in the number of late-coming parents in equilibrium. Mathematically, we have $x^* \ge x_4$ if $f_1 \le f < f_0$, where x^* denotes a stable equilibrium in our model.

Those parents whose types are located between x_1 and x_3 in Fig. 1 may be dubbed the "middle class," and it seems normal that they occupy the major portion of the parent population. As can be seen from Fig. 1, the extant portion of parents who come late is crucial

---

[16] The possibility of two stable equilibria in this kind of social norm model is clearly envisioned by Akerlof (1980). G&R explored the possibility of multiple equilibria in their game theory rather than in their social norm model. They suggested that the introduction of the fine might have shifted the game from an equilibrium with the low portion of late-coming parents to an equilibrium with the high portion. However, G&R did not specify explicitly the mechanism through which this shift occurs.
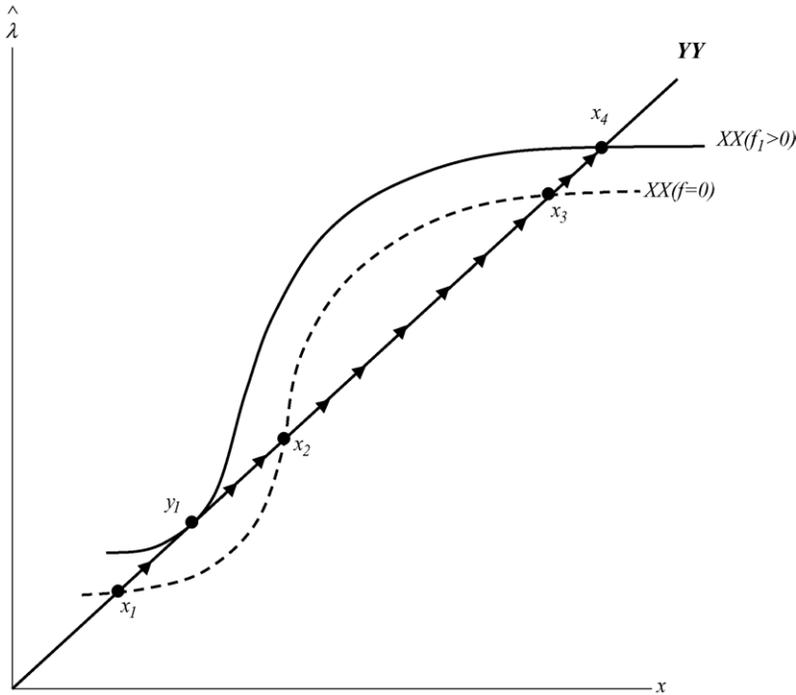
Fig. 2. A fine is imposed.

for the behavior of the middle class. When the extant portions of late-coming parents are lower than x_2, all parents in the middle class will choose to be on time in equilibrium, and therefore, the low late-coming equilibrium x_1 will be realized. By contrast, when the extant portions of late-coming parents are higher than x_2, all parents in the middle class will choose to come late in equilibrium, and as a result the high late-coming equilibrium x_3 will be realized. The key here is that the psychological cost inflicted on a late-coming parent is negatively related to the extant portion of late-coming parents (i.e. $c_x < 0$). This to a large extent explains why the psychological cost inflicted will be large enough to deter the middle class from coming late if the extant portions of late-coming parents are lower than x_2, but the opposite result will occur if the extant portions of late-coming parents are higher than x_2.

When a fine is introduced to shift the locus XX upward, the consequence for the middle class is that the regime associated with high psychological costs (those between x_1 and x_2) will contract, while the regime associated with low psychological costs (those between x_2 and x_3) will expand. As the fine reaches the threshold level f_1, the high psychological cost regime will be completely wiped out as shown in Fig. 2. This triggers a bandwagon effect such that parents in the middle class (those whose types are between y_1 and x_4) move en masse from being on time to being behind time. This is the reason for a significant increase in the number of late-coming parents as the fine is introduced.
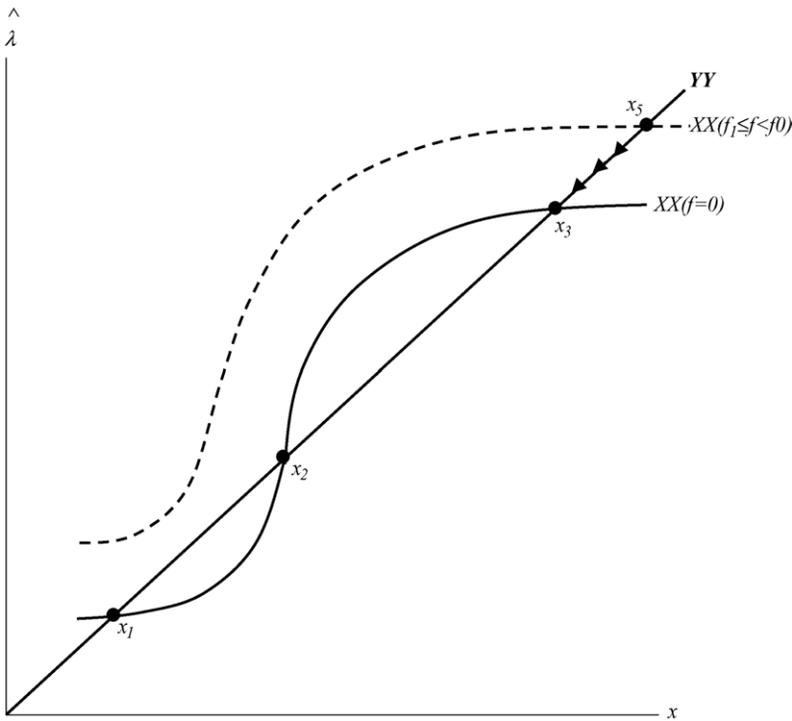
Fig. 3. The fine is removed.

4.3. The fine is removed

Suppose that the new equilibrium after the introduction of a monetary fine is at point x_5 as shown in Fig. 3. The locus *XX* that corresponds to this new equilibrium is represented by *XX* ($f_1 \leq f < f_0$) in the figure.

Now let us remove the fine imposed. The locus *XX* ($f_1 \leq f < f_0$) will then shift downward and return to the locus with no fine, that is, *XX* ($f = 0$). As shown in Fig. 3, however, the equilibrium after the removal of the fine will not return to point x_1 but will settle at point x_3. This result is simply due to the fact that the new status quo after the imposition of the fine is at point x_5. Thus, we have obtained the result where the number of late-coming parents remains higher and does not return to the pre-fine level after the fine is removed.[17]

The key here lies in there being two stable equilibria, x_1 and x_3, associated with the no-fine locus *XX* ($f = 0$). Which equilibrium will be realized critically depends on whether

---

[17] According to our model, the portion of late-coming parents will keep increasing (decreasing) after the introduction (removal) of the fine until it has reached the new equilibrium (see Figs. 2 and 3). However, the data collected in the G&R study move up and down after the introduction (removal) of the fine. A possible way of accounting for the deviation between the prediction of our model and the data is to add an error term to our theoretical model as is the case in regression analysis. After all, our model is a simplification of reality, abstracting from many details. A theoretical model like ours may capture the trend of the actual movement but it cannot possibly take into consideration all the up and down disturbances.

the status quo portion of late-coming parents is a low number or a high number. Intuitively speaking, the bite or effectiveness of the social norm against coming late hinges on the number of late-coming parents at the status quo. If coming late is widespread and rampant, the social norm will be too weak to generate a substantial psychological cost to inflict on late-coming parents. By contrast, if coming late is confined to a few parents only, the social norm will furnish a substantial psychological cost for late-coming parents. This explains why, depending on where the status quo portion of late-coming parents is, the same no-fine policy can result in very different portions of late-coming parents in equilibrium.

5. Discussion

As noted by Stigler (1987), the law of diminishing returns may set in from the beginning. If this is the case, the functional relationship between $\hat{\lambda}$ and x as represented by the locus *XX* in Fig. 1 will be replaced by the locus *XX′* in Fig. 4. As the arrows in Fig. 4 indicate, there are two stable equilibria, $x'_1 = 0$ and x'_3. If the status quo portion of late-coming parents
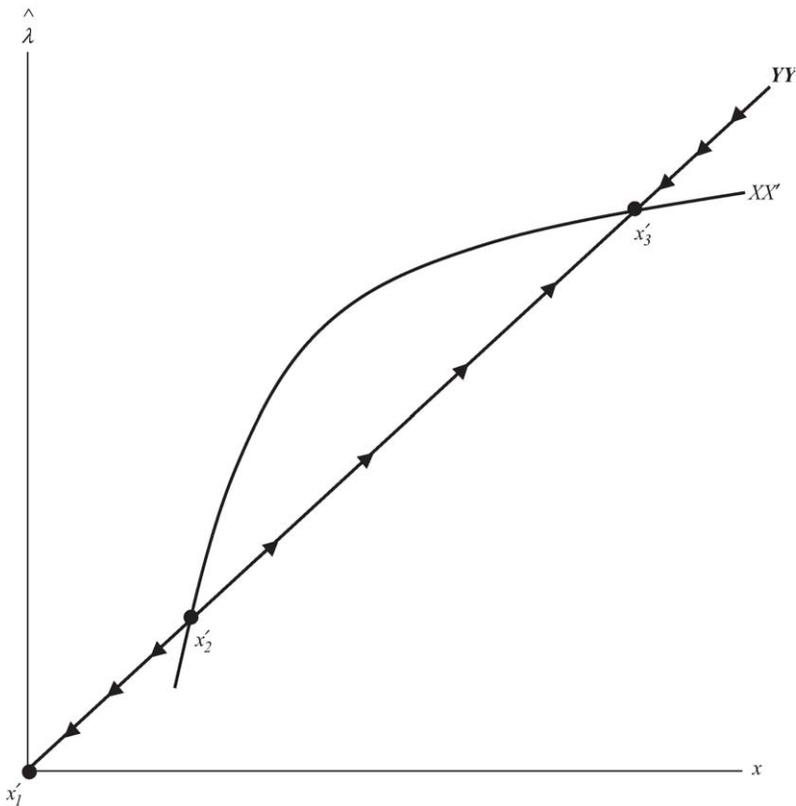

Fig. 4. Diminishing from the beginning.

is not large enough before the fine is introduced, then the initial equilibrium must occur at $x'_1 = 0$. However, this corner solution implies few or even no late-coming parents, which seems inconsistent with the data collected in the G&R field study.

It is more than likely that the psychological cost in question would be none-existent or next to none-existent if parents were to live in isolation. Through the incorporation of the C term in (1), individual decisions will be made in *social context* rather than in *isolation*. It is a game against your neighbors, not against nature or the authorities. This, we believe, is the major deviation of our model from the standard model in the deterrence hypothesis.

Our model is built on the "on time" social norm and the psychological cost associated with its violation. Although not mutually exclusive, the story of our model is clearly different from the G&R story of an incomplete contract and parents' changes in perception regarding the contract. In our view, G&R have tended to emphasize the contradiction of the deterrence hypothesis with their finding and have sought alternative, competing models to replace it. By contrast, the predictive power of the deterrence hypothesis is qualified but not lost according to our model. First, there are two terms rather than one term in (5a). Even if the deterrence effect is dominated, it still exists. Second, and more importantly, the deterrence effect may dominate the social norm-psychological cost effect so that the sign of (5a) becomes negative. This is possible if the fine imposed is relatively large (i.e. $f > f_0$).

When the sign of (5a) is negative, the locus XX ($f > 0$) will be located lower instead of higher than the locus XX ($f = 0$) in terms of our figures. Consider the locus XX ($f_2 > f_0$) in Fig. 5. The unique feature of this locus is that it is tangential to the locus YY at point y_2. It is clear from Fig. 5 that once the fine is increased to f_2 or higher, the portion of late-coming parents will be smaller than x_1 at the new equilibrium.[18] This result is interesting and somewhat paradoxical, since the "right" policy for the return to a low-level equilibrium is not the removal of the fine imposed, but a further increase in the fine (so that the new locus will be XX ($f_2 > f_0$) or lower, rather than XX ($f = 0$)). It is important to recognize here that we are not talking about imposing a very large fine so that no parents would dare to come late. The deterrence-dominating outcome may result in our model even if the fine imposed is smaller than the value of time (i.e. $\hat{\lambda}_f < 0$ may hold under $f < v$).

Our analysis above implies a novel prescription for the "stick" policy: fine enough or don't fine at all. A symmetric prescription for the "carrot" policy would then be: pay enough or don't pay at all. This is indeed the conclusion reached by Gneezy and Rustichini (2000b), who addressed the relationship between performance and material incentives. In their paper Gneezy and Rustichini conducted experiments, showing that for all positive but small enough compensations, the introduction of monetary compensation would worsen rather then improve performance, though a higher monetary compensation did produce a better performance once money was offered. In the sense that the effect of monetary incentives on desirable behavior is not monotonic, our result resembles this finding.

---

[18] Catastrophe theory is a mathematical theory that studies how a continuous variation in parameters can cause discontinuous effects or jumps in the large magnitude. The discontinuous effects or jumps are known as "catastrophes." From Figs. 2 and 5, we see that our model displays some feature of "catastrophes" in the sense that there are large, discontinuous changes in the *equilibrium* portion of late-coming parents as the fine imposed varies continuously. For an introduction to catastrophe theory and its applications in economics, see Rosser (2000).
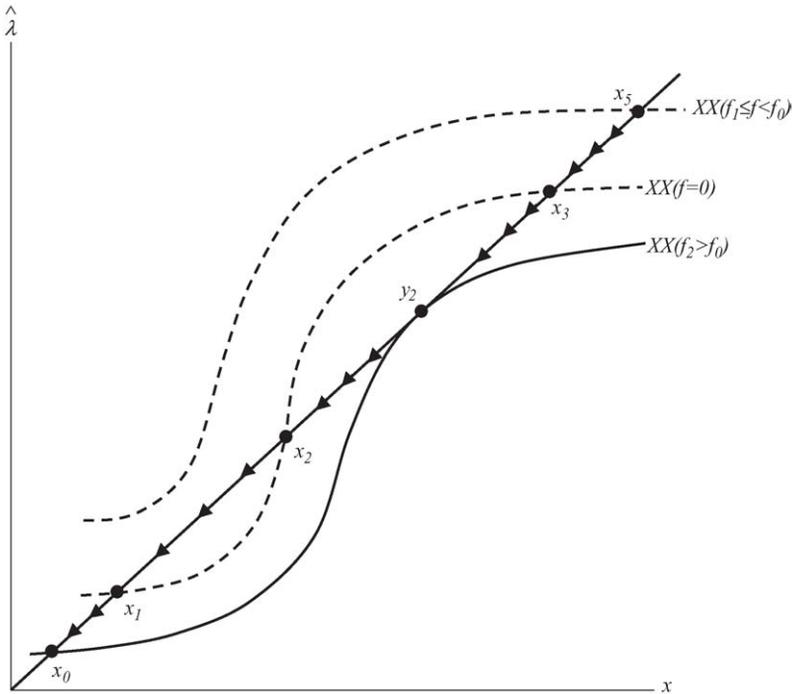
Fig. 5. The fine is increased further.

It is interesting to observe that in accounting for the G&R finding, our model as well as G&R's two models each rely on the relaxation or violation of the "all else equal" clause in the deterrence hypothesis. The introduction of a monetary fine changes parents' perceptions of the social situation in which they are involved in G&R's two models, while it erodes the effectiveness of the "on time" norm against late-coming parents in our model. G&R questioned the predictive strength of the deterrence hypothesis on the basis of the argument that the "all else equal" clause included in the deterrence hypothesis might be hard to satisfy in real life. In the light of G&R's two models and our model, the moral is perhaps that before applying the deterrence hypothesis to make a prediction, one should always check whether there are any induced side effects that may render the "all else equal" clause inapplicable and whether the induced side effects may more than offset the deterrence effect. In our model, for example, one must know whether the elasticity ε is significant enough so as to uphold the inequality (6).

6. Concluding remarks

G&R (p. 15) emphasized that the results from their field study "are probably more than a curious finding." Several recent empirical studies collaborate G&R's emphasis, showing that punishment may backfire and enhance rather than inhibit "crimes." Sheffrin and Triest

(1992) examined people's attitudes on tax compliance and found that tax audits may negatively affect attitudes and create an adversarial relationship between the taxpayer and the tax agency. Bewley (1999, Chapter 8) reported that the company managers he interviewed believed that threats of pay cuts or of dismissal are counterproductive since they imply coercion. Fehr and Gächter (2002) showed in their experiments that incentive contracts involving explicit punishment may undermine voluntary cooperation. Bohnet et al. (2001) studied the enforceability of incomplete contracts and demonstrated experimentally that increasing the probability of enforcement may not raise the likelihood of contract performance; however, similar to our finding, higher performance will eventually result when levels of enforcement are set high enough.[19]

Besides the finding that punishment may backfire, another major result found in G&R is that "crimes" committed may persist or display inertia once they pass some threshold and reach a high level. Gaviria (2000) documented a similar phenomenon for the violent crime in Colombia. The Colombian violent crime escalated rapidly during the 1980s. Similar to our model, Gaviria rationalized this escalation as a movement from a low- to a high-crime equilibrium and found some empirical evidence consistent with his rationalization. Once a high-crime equilibrium is reached, Gaviria argued on the basis of a wealth of casual evidence that crime becomes the way of life in Colombia and it even receives some legitimacy in the eyes of many Columbians.[20]

The erosion of social norms via interactions across "criminals" plays a crucial role in our explanation of G&R's second result. This mechanism may well explain other phenomena. U.S. consumer bankruptcy filing rates tripled from 1984 to 1991. This run-up in filing rates constitutes one of the greatest puzzles in American bankruptcy law since both legal changes and economic conditions would predict otherwise. Buckley and Brinig (1998) sought to account for the puzzle and argued that certain social norms, as captured by the strength of social networks, conservative and hierarchical attitudes, and the social stigma of promise-breaking, may have had an impact on the consumer's decision to file bankruptcy. Their empirical result suggests that, more than anything, the explosion in filing rates is attributable to a shift in social norms. Sampson et al. (1997) interviewed nearly 8800 residents in 343 neighborhoods and asked them whether in their neighborhood people can be trusted, share the same values, get along with one another, and so on. They employed the responses to construct an index of "collective efficacy" and found that this index can explain a large proportion of crime variation across neighborhoods. Consistent with our mechanism, Sheffrin

---

[19] Incentives may take the negative form of punishment, but they may also take the positive form of reward. In his influential book *The Gift Relationship*, Titmuss (1970) advanced the argument that payments for donations of human blood may cause altruism to decline and negatively affect people's willingness to donate blood. Stewart (1992) presented a theoretical model of blood supply *a la* Akerlof (1980) to illustrate Titmuss's argument. An argument analogous to Titmuss's appears in the field of social psychology. It is the so-called "hidden costs of reward": externally imposed material reward may undermine an individual's intrinsic motivation to carry out a pleasant activity. Professor Frey and his co-authors have extensively explored the relevance of this literature to economics; see Frey (1994), Frey et al. (1996), and Frey and Oberholzer-Gee (1997).

[20] There is the so-called "overload theory" in the crime and punishment literature: crime will increase still further if funding for law enforcement does not increase enough to prevent the expected penalty from falling. This theory can also generate multiple equilibria, allowing for the possibility that an economy will move from a low- to a high-crime equilibrium. See Rasmusen (1996) for the relevant references.

and Triest showed empirically that publicity regarding the existing degree of tax evasion may increase evasion further as taxpayers alter their perceptions of social norms. Falk and Fischbacher (2002) detected in their controlled laboratory experiments that subjects have a lower propensity to comply with the norm that "they should not steal" the more that others steal.

We conclude our paper by highlighting several salient features that are displayed in our presented figures. These features summarize the gist of our formal model.[21]

1. *Multiple equilibria*: Depending on initial conditions, seemingly identical situations may result in very different outcomes (for example, the equilibrium portions of late-coming parents, x_1 and x_3, are both identified with no fine, $f = 0$; see Fig. 1).
2. *Sudden jumps*: Aggregate behavior may jump discontinuously even though parameters within the economy vary smoothly (for example, the equilibrium portion of late-coming parents will jump from y_1 to x_4 when the fine imposed reaches f_1; see Fig. 2).
3. *No middle ground*: The level of aggregate behavior is either low or high and there is no middle ground in between (for example, the portion of late-coming parents, x_2, is an unstable equilibrium and hence realized equilibria will be either low (x_1) or high (x_3); see Fig. 1).
4. *Cycles*: Aggregate behavior may exhibit cyclical patterns as parameters within the economy vary (for example, the equilibrium portion of late-coming parents will revert back to a low level x_0 when the fine is raised to f_2; see Fig. 5).

To generate some or all of these features, the modeling strategy is actually simple. First of all, consider a binary-choice decision problem[22]: whether to commit a crime, to evade tax, to dine at a restaurant, to adopt a PC operating system, to shirk in the workplace, and to participate in a political protest, and so on and so forth. Next, let the individual decision hinge on how many others have made the same choice so that there exists the "snowballing" effect of social interactions ($\hat{\lambda}_x > 0$ in our model). This hinge may be due to the bite of social norms as in our model or some other reasons (say, the pleasure from consuming a good is greater when many people consume it as suggested in Becker, 1991). The third step is to vary parameters within the economy so as to shift the locus upward or downward in a way similar to shifting the *XX* curve in our model. The resulting features obtained may explain a variety of phenomena, including the frequently observed cyclical pattern in crime rates and their dramatic changes over time and space,[23] the possible avalanche in people's tax evasion,[24] fickle consumer demand,[25] the mass switching of operating systems from

---

[21] These features are associated with some elementary catastrophes; see Rosser (2000) for the detail.

[22] A continuous-choice decision problem can be reduced to a binary-choice one. For example, consider the decision about tax evasion. Once the decision to evade tax has been made, there will be an optimal amount of tax evaded. We can then confine the analysis to comparing the payoff from no evasion with that from optimal evasion.

[23] See, for example, Glaeser et al. and the references therein for the evidence.

[24] Myles and Naylor (1996) considered a tax-evasion model with group conformity and social customs. They obtained a related result: an economy previously characterized by no evasion may be destroyed by a small change in the tax rate, resulting in a potential epidemic of tax evasion.

[25] See Becker.

DOS to Windows,[26] employers' hesitation to cut worker pay even during the recession,[27] and the tipping of political protests over into revolution.[28] It is hoped that these illustrations may be persuasive of potentially wide-ranging applications of our formal model.

References

Adams, S.J., 1963. Toward an understanding of inequality. Journal of Abnormal and Social Psychology 67, 422–436.

Akerlof, G.A., 1980. A theory of social custom, of which unemployment may be one consequence. Quarterly Journal of Economics 94, 749–775.

Akerlof, G.A., Yellen, J.L., 1990. The fair wage-effort hypothesis and unemployment. Quarterly Journal of Economics 105, 255–283.

Akerlof, G.A., Yellen, J.L., 1991. How large are the losses from rule-of thumb behavior in models of the business cycle? In: Nordhaus, W., Watts, H.W. (Eds.), Money, Macroeconomics, and Economic Policy: Essays in Honor of James Tobin. MIT Press, Cambridge, pp. 59–78.

Becker, G.S., 1991. A note on restaurant pricing and other examples of social influences on price. Journal of Political Economy 99, 1109–1116.

Bennett, W.J., 1993. The Book of Virtues: A Treasury of Great Moral Stories. Simon & Schuster, New York.

Bewley, T.F., 1999. Why Wages Don't Fall During a Recession. Harvard University Press, Cambridge.

Bohnet, I., Frey, B.S., Huck, S., 2001. More order with less law: on contract enforcement, trust, and crowding. American Political Science Review 95, 131–144.

Booth, A.L., 1985. The free rider problem and a social custom model of trade union membership. Quarterly Journal of Economics 100, 253–261.

Buckley, F.H., Brinig, M.F., 1998. The bankruptcy puzzle. Journal of Legal Studies 27, 187–207.

Cooter, R., Ulen, T., 2000. Law and Economics, third ed. Addison-Wesley, Massachusetts.

Falk, A., Fischbacher, U., 2002. Crime in the lab-detecting social interaction. European Economic Review 46, 859–869.

Fehr, E., Falk, A., 2002. Psychological foundation of incentives. European Economic Review 46, 687–724.

Fehr, E., Gächter, S., 2002. Do incentive contracts undermine voluntary cooperation. Working Paper No. 34, Institute for Empirical Research in Economics, University of Zurich.

Frey, B.S., 1994. How intrinsic motivation is crowded out and in. Rationality and Society 6, 334–352.

Frey, B.S., Oberholzer-Gee, F., 1997. The cost of price incentives: an empirical analysis of motivation crowding-out. American Economic Review 87, 746–755.

Frey, B.S., Oberholzer-Gee, F., Eichenberger, R., 1996. The old lady visits your backyard: a tale of morals and markets. Journal of Political Economy 104, 1297–1313.

Gaviria, A., 2000. Increasing returns and the evolution of violent crime: the case of Colombia. Journal of Development Economics 61, 1–25.

Glaeser, E.L., Sacerdote, B., Sheinkman, J.A., 1996. Crime and social interactions. Quarterly Journal of Economics 111, 507–548.

Gneezy, U., Rustichini, A., 2000a. A fine is a price. Journal of Legal Studies 29, 1–17.

Gneezy, U., Rustichini, A., 2000b. Pay enough or don't pay at all. Quarterly Journal of Economics 115, 791–810.

Kuran, T., 1991. Now out of never: the element of surprise in the Eastern European revolution of 1989. World Politics 44, 7–48.

Lange, R., McDade, S., Oliva, A., 2001. Technological choice and network externalities: a catastrophe model analysis of firm software adoption for competing operating systems. Structural Change and Economic Dynamics 12, 29–57.

---

[26] See Lange et al. (2001) for an empirical study.

[27] According to the views of business people interviewed by Bewley, pay cuts are likely to trigger widespread shirking among workers and cause a large fall in labor productivity.

[28] For a vivid account of the sudden collapse of Eastern Europe's communist regimes, see Kuran (1991).

Lindbeck, A., Nyberg, S., Weibull, J.W., 1999. Social norms and economic incentives in the welfare state. Quarterly Journal of Economics 114, 1–35.

Mundell, R.A., 1968. Man and Economics. McGraw-Hill, New York.

Myles, G.D., Naylor, R.A., 1996. A model of tax evasion with group conformity and social customs. European Journal of Political Economy 12, 49–66.

Naylor, R., 1989. Strikes, free riders, and social customs. Quarterly Journal of Economics 104, 771–785.

Posner, R.A., Rasmusen, E.B., 1999. Creating and enforcing norms, with special reference to sanctions. International Review of Law and Economics 19, 369–382.

Rasmusen, E., 1996. Stigma and self-fulfilling expectations of criminality. Journal of Law and Economics 39, 519–543.

Romer, D., 1984. The theory of social custom: a modification and some extensions. Quarterly Journal of Economics 98, 717–727.

Rosser Jr., B.J., 2000. From Catastrophe to Chaos: A General Theory of Economic Discontinuities, second ed. Kluwer Academic Press, Boston.

Sampson, R.J., Raudenbush, E.F., 1997. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 277, 918–924.

Sheffrin, S.M., Triest, R.K., 1992. Can brute deterrence backfire? Perceptions and attitudes in taxpayer compliance. In: Slemrod, J. (Ed.), Why People Pay Taxes: Tax Compliance and Enforcement. University of Michigan Press, Ann Abor, pp. 193–218.

Stewart, H., 1992. Rationality and the market for human blood. Journal of Economic Behavior and Organization 19, 125–143.

Stigler, G.J., 1987. The Theory of Price, fourth ed. Macmillan, New York.

Titmuss, R., 1970. The Gift Relationship: From Human Blood to Social Policy. George Allen and Unwin, London.