האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

THE ECONOMICS OF SHAME:
WHY MORE SHAMING MAY
DETER LESS

by

ALON HAREL and ALON KLEMENT

מרכז לחקר הרציונליות

CENTER FOR THE STUDY
OF RATIONALITY

THE ECONOMICS OF SHAME: WHY MORE SHAMING MAY DETER LESS[†]

ALON HAREL[*] & ALON KLEMENT[**]
August 24, 2005

ABSTRACT

This paper investigates the effectiveness of shaming penalties. It establishes that there may be an inverse relation between the rate of shaming penalties and their deterrent effects - the more people are shamed the less deterring shaming penalties become. This conclusion is based on a *search model* in which the costs of searching for law-abiding partners increase with the rate of shaming, and may lead to lower expected sanction for offenders. The inverse relation between the rate of shaming penalties and their effectiveness is later used to show that increasing the probability of detection, increasing the magnitude of shaming penalties or reducing the number of wrongful acquittals does not necessarily increase the deterrent effects of shaming penalties (and may, in fact, decrease these effects).

[*] Philip P. Mizock & Estelle Mizock Professor of Law, Faculty of Law, Hebrew University, Jerusalem. msalon@mscc.huji.ac.il

[**] Radziner School of Law, Interdisciplinary Center, Herzliya. klement@idc.ac.il

1. Introduction

Shaming penalties operate primarily by publicizing an offender's illegal conduct in a way which enhances sentiments of disapproval, disgust and contempt. They expose the wrongdoer, inflict psychological and social costs on her by raising feelings of guilt and remorse and shunning her from her community, impose economic costs by limiting her ability to conduct commercial interactions with others, and reinforce the prevailing social norms that disapprove of such behavior.[1] Such penalties differ from traditional legal sanctions such as imprisonment or fines in that their deterrent effects depend largely on a network of mutual social understandings; in particular, their effectiveness depends on the reaction of members in society or in the market to the public information concerning the offender's behavior. The literature is filled with debates concerning the justifiability and effectiveness of shaming penalties,[2] as well as numerous discussions concerning their very legality.[3]

This article points out some limitations of shaming penalties that have gone unaccounted for – limitations that constrain their extensive use. More particularly, it is argued that wide-ranging use of shaming penalties is likely to erode their effectiveness, and that their extensive use as a substitute for traditional sanctions may

---

[1] See *Note: Shame, Stigma, and Crime: Evaluating the Efficacy of Shaming Sanctions in Criminal Law* 116 HARVARD L.REV. 2187 (2003).

[2] See, e.g., Dan Kahan, *What do Alternative Sanctions Mean?* 63 U. Chi. L. Rev. 591, 592 (1996) (advocating the use of shaming penalties as a substitute for traditional sanctions such as incarceration and fines); Dan M. Kahan, & Eric A. Posner, *Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines* XLII JOURNAL OF L.&ECON. 365, 367-68 (1999) (advocating the use of shame penalties for white collar crimes).

[3] Some courts found certain shaming penalties to be unconstitutional. See, e.g., *People v. Meyer*, 680 N.E.2d 315, 315-16 (Ill. 1997) (overruling the decision of the Illinois trial court to order the defendant to erect at his home a four foot by eight foot sign with eight inch high lettering that reads "Warning! A Violent Felon Lives Here. Enter at your own Risk"); *People v. Letterlough*, 613 N.Y.S. 2d 687, 688 (App. Div. 1994) (overruling the decision of the lower court to require the defendant to affix to the license plate of any vehicle he drove a sign stating "convicted DWI"). For a discussion and a critique of the appellate courts decisions to strike down shaming, see Aaron S. Book, *Note: Shame on You: An Analysis of Modern Shame Punishment as an Alternative to Incarceration* 40 WM AND MARY L. REV. 653, 670-72 (1999). Other courts have ruled that shaming penalties are constitutional. See, e.g., *United States v. Gementera*, 379 F. 3d 596 (9th Cir. 2004) (upholding the decision of to require a convict to wear a signboard proclaiming his guilt); *Goldschmitt v. State*, No. 69,094, Supreme Court of Florida, 496 So. 2d 142; (upholding the decision of the trial court to require a defendant to place a sticker read: "CONVICTED DUI – RESTRICTED LICENSE"); United States v. Coenen, 135 F.3d 938 (upholding the decision requiring the defendant to publish notice in the official journal of the governing authority of the parish; United States v. Schecter, 13 F.3d 1117 (7th Cir. 1994) (upholding the decision to require the defendant to notify all future employers of the defendant past criminal conduct (tax evasions); *Ballenger v. State*, 436 SE2d 793 (Ga. Ct. App. 1993) (upholding imposition of a condition requiring a probationer to wear a fluorescent pink plastic bracelet imprinted with the words "DUI CONVICTED").

undermine their deterrent effects. The effectiveness of shaming penalties is inversely related to the rate of shaming in the society. [4] The more people are shamed, the less effective shaming penalties may become.[5]

To demonstrate the inverse relations between the rate of imposing shaming penalties and the effectiveness of shaming penalties we use a market search model in which the search costs of law-abiding citizens (resulting from shunning the shamed and searching for law abiding commercial partners) increase with the size of the shamed population. We show that shaming penalties are often effective only when they are rarely used; the more they are used, the less effective they may become.

The inverse relations between the number of shamed and the deterrent effect of shaming penalties may translate into three distinctive and counter-intuitive features that ought to influence the ways in which these penalties are used. We show that increasing the *probability* of detection and conviction, increasing the *magnitude* of shaming penalties, and increasing the *accuracy* of the legal process, i.e., reducing the number of legal mistakes, may each dilute the deterrent effects of shaming penalties. These three effects ought to be considered while determining the optimal investment in detection, the optimal size of the shaming sanctions and the optimal investment in judicial accuracy. Furthermore, theses effects should alert policy makers that the maximal level of deterrence feasible through shaming might prove to be much lower than the one possible when using traditional sanctions.

---

[4] Some previous accounts raised a similar conjecture. Thus for instance Tony Massaro stated that: "But…if shaming penalties were imposed equally on all offenders who commit similar offences, this could undercut the impact of these penalties; the more people subject to shaming, the less it compromises one's social status – it could even begin to elevate it in some cases (If, e.g., five cars in the neighborhood bear 'DUI' plates, then the plates may lose some ego ideal and social status shattering effect.) Just as jail time has lost its stigma within certain subcultures, so might pillory time lose its sting if many members of the subculture have endured it." See Toni M. Massaro, *The Meanings of Shame: Implications for Legal Reform* 3 PSYCH. PUB. POL. & L. 645, 697-98. Yet, Massaro's analysis is based on psychological conjectures. Instead, our analysis is founded on the analysis of rational behavior in social groups and in the market. It is rationality of individuals (both of law-abiding citizens and that of offenders) which dictates that shaming penalties lose their effectiveness if used extensively.

[5] The first to model shaming in a rational expectations model was Eric Rasmusen, *Stigma and Self Fulfilling Expectations of Criminality*, 29 J. LAW & ECON. 519 (1996). For reasons we explain below, Rasmusen found that increasing the rate of shaming would enhance deterrence, in contrast to our conclusions. In a paper based on Rasmusen's model Patricia Funk, *On the Effective Use of Stigma as Crime Deterrent*, 48 EUROPEAN ECON. REV. 715 (2004) shows that shaming has little effect on stigmatized ex-convicts and consequently increasing the rate of shaming may have negative effects on deterrence (resulting from the fact that ex-convicts are not deterred by shaming). Yet, this effect is attributable exclusively to the effects it has on ex-convicts. This, again, stands in contrast to our findings. Finally, Lawrence Blume, *Stigma and Social Control*, unpublished manuscript available at http://ideas.repec.org/p/wpa/wuwpga/0312002.html analyzes a dynamic population model and explores the evolution of stigma and crime. It assumes that stigma is inversely correlated with the rate of shamed individuals, an assumption which we demonstrate as a result in our model.

Section 2 analyzes the costs borne by individuals and the effects that these costs may have on their willingness to cooperate with the shaming scheme. It also demonstrates that the willingness to bear the costs of shaming differs between different groups of shamers. Section 3 establishes our main claim - that the deterrent effects of shaming may be inversely related to the number of individuals who are shamed. It also demonstrates that an increase in the probability of detection, in the size of the sanction or in the accuracy of the judicial process may result in a decrease in the deterrent effects of shaming penalties. Section 4 briefly demonstrates the relevance of the search model to contemporary shaming penalties legislation. Section 5 concludes.

2. Why do People Shame?

Three explanations for the effectiveness of shaming penalties are often provided in the literature.[6] First, it is argued that shaming generates guilt and remorse on the part of the offender;[7] second, it reinforces respect towards legal norms and entrenches a law abiding culture;[8] and third, shaming imposes costs on offenders by identifying them and disseminating information which generates social and professional isolation and alienation from law abiding society. Individuals are deterred because other individuals, law abiding individuals in particular, would limit their social or professional interaction with them as a result of their being subjected to a shaming penalty.[9] It is this latter mode of operation that we examine in this paper.[10]

---

[6] For a survey of these three effects, see Note, *supra* note 1.

[7] Common sense suggests that genuine remorse is triggered by committing the crime, not by publicizing it. Yet, some psychologists argue otherwise and regards publicity as a means of triggering guilt and remorse. See Richard H. Smith & J. Matthew Webster, W. Gerrod Parrot & Heidi L. Eyre, *The Role of Public Exposure in Moral and Nonmoral Shame and Guilt*, 83 JOURNAL OF PERSONALITY & SOCIAL PSYCHOLOGY, 138 (2002) ("the linking of public exposure with shame is unmistakable.").

[8] Shaming penalties help generating or at least strengthening and entrenching respect towards the law, i.e., help in forming a law-abiding culture. Under this view, shaming penalties fortify and reinforce legal norms through mechanisms of preference formation and "belief-dependent propensities" to obey the law. See, supra note Kahan & Posner, *supra* note 2 376-380.

[9] See Kahan, *supra* note 2 at 638 ("The consequences of shaming penalties are extremely unpleasant. Those who lose the respect of their peers often suffer a crippling diminishment of self-esteem. Moreover, criminal offenders are as likely to be shunned in the marketplace as they are in the public square, leading to serious financial hardship."); Stephen P. Garvey, *Can Shaming Punishments Educate?* 65 U. CHI. L. REV. 733, 752 (1998) (noting that the person who is subject to shame penalties "may suffer adverse consequences from members of the community, who may…refuse to engage in various forms of social and economic intercourse with him.") .

[10] Arguably the dissemination of information concerning a person's criminal record is also facilitated by the imposition of traditional sanctions such as incarceration and fines. The criminal trial itself

Shaming, under this explanation, presupposes the active cooperation of private individuals – individuals who have to invest resources in facilitating the shaming scheme.[11] Such a costly cooperation is needed because shaming penalties do not operate directly by imposing "public sanctions" (e.g. fines or imprisonment) on offenders but indirectly by facilitating the imposition of "private sanctions" on offenders, as well as by reinforcing the disposition to impose such private sanctions. The effectiveness of private sanctions is based on the active cooperation of "private enforcers" and their willingness to invest resources in imposing private sanctions.

Other things being equal, the more resources private enforcers have to invest in the imposition of private sanctions the lesser the willingness of private enforcers to cooperate with the shaming scheme and consequently the less effective shaming penalties become. In order to evaluate the effectiveness of shaming penalties, it is therefore crucial to identify the costs that private enforcers have to incur in order to cooperate with the scheme of shaming.

Private enforcers are not a homogenous group. There are different types of private enforcers and the willingness of private enforcers to invest in cooperating with the shaming scheme differs from one group of enforcers to another. Individuals who cooperate with the shaming scheme can be usefully classified into two groups. Each one of these groups faces a different set of incentives and consequently reacts differently to an increase in the probability of detection and shaming. The first group – the *offenders-centered enforcers* – consists of individuals who invest in private sanctions which are directed **against offenders**. The second group – the *shamed-centered enforcers* -- consists of individuals who target not offenders but **shamed individuals**.[12]

(irrespective of what the sanctions are) is public and the criminal record of a person is often available. Why are therefore shaming penalties particularly distinctive in this respect? Why are their informational consequences perceived to be particularly damaging to the criminal? The answer is that shaming penalties are especially designed to facilitate easy and cheap dissemination of information and thereby result in a much broader dissemination of information than traditional sanctions. Publicity is often the primary and sometimes the exclusive component of shaming penalties. Moreover, the theatrical exposure of the criminal, his evil deeds and sometimes his (forced) apology attracts broad public attention and consequently the relevant information is likely to be more broadly disseminated. For a discussion of the special communicative effects of different sanctions, see Eric Posner, LAW AND SOCIAL NORMS 108-110 (2000).

[11] See, e.g., Posner, supra note 10 at 90 ("The signal could be costly either because this individual values his interactions with the bad type, or because he discounts the future highly"); Kahan & Posner, supra note 2 at 372 ("[S]haming is not as cheap as it seems" since "shaming involves all kinds of hidden costs incurred by people who shun the offender")

[12] While the precise distinction drawn by us has not been made in the literature some advocates of shaming penalties have drawn attention to different categories of shamers which constitute special

Offenders-centered enforcers refrain from interacting with individuals who have actually committed the offence (irrespective of whether they are shamed or not). Their utility from interaction with a person (be it social or business interaction) is inversely correlated with the commission of an illegal act by that person. An interaction with an individual who committed the illegal act is less beneficial (or more costly) for an offenders-centered enforcer.

People may negatively react to the past commission of an offence when the commission of the offence is indicative that the goods or services that are likely to be provided by an offender are inferior to those that are likely to be provided by non-offenders. For example, people are less likely to hire a cab-driver who has driven before while being intoxicated, because driving with such a person is risky. Parents would typically not hire a babysitter who has committed sex-offences because of their belief that such behavior indicates dispositions to commit similar offences in the future. In both cases the services that are likely to be provided by offenders are of lesser quality or involve greater risks than the services provided by non-offenders.

Shamed-centered enforcers refrain from interacting with individuals who are publicly and officially identified and labeled as offenders. They do not care whether the individuals they interact with are offenders or not. They are, however, reluctant to interact with shamed individuals. Such reluctance may be attributed to the unwillingness to be publicly observed interacting with shamed individuals. Interaction with the shamed might signal to third parties that those interacting with them are also "bad types". This common sense wisdom is often articulated by pieces of folk wisdom such as "tell me who your friends are, and I will tell you who you are", or "birds of the same feather flock together", or "lay down with dogs wake up with fleas." In contrast, refraining from interacting with the shamed may signal to third parties that those who refrain are good types who share their moral outlook and uphold moral values.

---

cases of the categories of offenders-centered and shamed-centered individuals. Thus, for instance, Kahan & Posner argued as follows:

"People avoid the offender for two reasons (1) the offender has been revealed as a bad type, who is thus likely to be unreliable in cooperative endeavors; and (2) even to the extent that it might be profitable to continue to deal with the offender (because he has special skills, for example), by ostentatiously avoiding him, one shows that one belongs to the good type and thus reveals oneself to be an attractive partner to others." See Kahan & Posner, *supra* note 2.

The first category described by Kahan & Posner is a special case of an offender-based enforcer while the second category is a special case of a shamed-centered enforcer.

Evidently, the preference not to interact with offenders and the preference not to interact with the shamed do not preclude each other and they can be present in the same person at the same time. A firm or a person may wish to shun offenders, yet their preference for doing so might be even more intense if these offenders have been publicly exposed and shamed. For the sake of expository clarity we nevertheless separate the two types in the following discussion.

Shaming penalties are clearly relevant for both offenders-centered and for shamed-centered enforcers. The offenders-centered enforcers use shaming as a proxy for identifying offenders even if their real intended targets are offenders rather than shamed individuals. In contrast shamed-centered enforcers do not use shaming merely as a proxy; their primary targets are those who have been publicly identified and labeled as offenders, i.e., the shamed. Hence, both groups would differentiate between shamed and non-shamed individuals and both types of enforcers would refrain from interacting with the shamed.

Yet, there is a major factor differentiating the behavior of offenders-centered and shamed-centered individuals. While the willingness of the offenders-centered and shamed-centered individuals to interact with the shamed is similar, it is their willingness to interact with the **non-shamed** that is different. More particularly, offenders-centered individuals bear costs when they interact with the **non-shamed** – costs that are not borne by shamed-centered individuals.

Shaming penalties are not imposed on all offenders. All private enforcers know that some offenders are not shamed despite having committed the offence. Yet, not all private enforcers react in a similar fashion to the imperfections of the enforcement system. The willingness of offenders-centered enforcers to interact with a non-shamed individual is affected by the possibility that that person may after all be a non-shamed offender. The potential gains that one may derive from the interaction with a non-shamed individual are discounted given the non-negligible probability that the non-shamed individual is, in reality, an offender. The possibility that the cab-driver one hires is an alcoholic is a matter of concern to a potential customer irrespective of whether the person has a DUI sign. In contrast, the shamed-centered enforcer is indifferent as to whether the person has committed an offence or not. His only concern is to identify the shamed individuals and to avoid interacting with them.

This difference between offenders-centered and shamed-centered enforcers has important unexpected ramifications. In particular, offenders-centered and shamed-

centered individuals react very differently to an increase in the probability of detection. An increase in the probability of detection increases the expected search costs of both offenders-centered and shamed-centered individuals, as they would have to search more until they find someone who was not shamed. At the same time, increasing the probability of detection is beneficial for offenders-centered enforcers because it allows them to infer more reliably the innocence of a person from the fact that that person has not been shamed and consequently it decreases the costs of interacting with the non-shamed.[13] For reasons that are elaborated in our model, the striking result in equilibrium is that the behavior of offender-centered individuals is unaffected by the rate of shaming. Those who prefer to search for a non-shamed individual continue doing so, and those who ignore shaming and interact both with the shamed and the non-shamed also do not change their behavior, irrespective of the rate of shaming.

Compare these effects to the effects that increasing the probability of detection has on the shamed-centered enforcers. Like offenders-centered enforcers, shamed-centered enforcers also face greater search costs resulting from an increase in the probability of detection. An increase in the rate of detection and shaming increases inevitably the search costs of all private enforcers. But, in contrast to the offenders-centered enforcers, the greater accuracy and reliability of the shaming scheme in detecting criminals does not generate any benefit for the shamed-centered enforcers because their benefits and costs from interacting with the non-shamed is unaffected. As a result, if the rate of shaming is increased, a shamed-centered enforcer who was indifferent between searching for a non-shamed and interacting with the shamed would now strictly prefer not to search. The rate of enforcement would therefore decline.[14]

To sum up, shaming penalties often operate by limiting and constraining both social and commercial opportunities. These constraints are the byproduct of the willingness of individuals – private enforcers - to "punish" the perpetrators of crime by limiting their social or professional interactions with them. Yet, this willingness on the part of individuals to punish the shamed may be costly, and its costs depend upon the number of shamed individuals. Furthermore, private enforcers' costs and benefits

---

[13] For a similar argument, see Rasmusen, *supra* note 5.

[14] A similar conjecture was raised by Kahan & Posner, supra note 2 at 372 ("If a lot of people are caught and shamed, then we might have no choice about whether to cooperate with him.")

would depend on their motivation for shunning the shamed. In particular, offenders-centered enforcers would benefit from the greater accuracy a higher rate of detection and shaming provides. Shamed-centered enforcers, on the other hand, would not enjoy such benefits. Since increasing the rate of detection increases the expected costs of search for non-shamed individuals, these differences would translate into different levels of private enforcement. As the following model demonstrates, the level of enforcement by offenders-centered individuals would not depend on the rate of shaming, whereas the level of enforcement among shamed-centered individuals would decline if the rate of detection and shaming increases.

3. Why More Shaming May Deter Less

A. The Search Model

Assume two disjoint sets of risk neutral individuals, *Sellers* and *Buyers*. The ratio between the number of sellers and the number of buyers is $r<<1$. Sellers and buyers play the following two stage game. In stage 1, each seller contemplates whether to commit an illegal act. Seller *i*'s utility from committing the illegal act, u_i, is uniformly distributed on $[0,1]$.[15] Only seller *i* knows his utility and whether he committed the illegal act. Each seller is audited by the state with probability p. If a seller is found to have committed the illegal act then this is made publicly known. Thus, at the end of stage 1 the seller population is divided into sellers who were announced *liable*, and sellers who were not (they are *non-liable*).

In stage 2 buyers are searching for some service which only sellers can perform. Each buyer has a unit demand for that service, whose price is fixed at $v<1$ (v 0).[16] A buyer's net benefit (net of v) from a service performed by seller *i* is equal to $B(u_i)$, where $B'(u_i)$ 0. If the buyer hires a liable seller he also bears a "shaming cost", denoted $S \geq 0$. The shaming cost is imposed on the buyer by affecting *her*

---

[15] The uniform distribution assumption entails no loss of generality. Suppose u is drawn from a strictly increasing continuous distribution function $F(\cdot)$ on $[\underline{u}, \overline{u}]$, where $\underline{u}$ 0<$\overline{u}$. It can be verified that if we take $F(\cdot)$ itself to be a random variable, it is uniformly distributed on $[0,1]$. Since $F(\cdot)$ is strictly increasing, any distribution of u can be recovered from the uniform distribution of $F(\cdot)$ using a monotonous transformation. See George Casella & Roger L. Berger, **STATISTICAL INFERENCE** 52, Theorem 2.1.4 (1990).

[16] Under this assumption there are sellers who would commit the illegal act even if they were certain to be caught. We make this assumption only to assure an "internal" equilibrium - that is an equilibrium in which some, but not all, individuals commit the illegal act. See note 20 below.

opportunities for business and social relationships.[17] It is assumed that $B(1) > S$, so the buyer derives an expected net positive benefit even if the service is provided by a (worst) type "1" seller. This assumption is made only for convenience of exposition and it is easily verified that our results carry through if it is relaxed.

Stage 2 is a multiple period stage. In every period each of the buyers picks a random seller and decides whether to hire him to perform the service or not. If and only if she does not hire that seller, she continues her search and picks another seller in the next period. Search costs vary among buyers. Buyer j's search costs are k_j. Sellers only know that k_j is drawn from a continuous distribution function $G(\cdot)$ on $[0,K]$ for some $K > S$.[18] We assume that production costs of the service are 0, so each seller's net utility from providing the service is v. Finally, each seller may provide the service to as many buyers as would want to hire him for it.

An outcome of this two stage game is given by the decision of each seller whether to commit the illegal act, and the decision of each buyer whether to search in each period for a non-liable seller.

Few points are worth notice. First, we assume a fixed and equal price for liable and non-liable sellers, above their reservation value. This would be the case if buyers have all the bargaining power, yet the price is bounded from below, above sellers' reservation value. Such an assumption seems appropriate for social contexts, where the service provided entails a social interaction, without any feasible pricing mechanism. It would also be plausible in business contexts that feature price rigidity (e.g. employment subject to an effective minimum wage control or subject to other forms of price regulations).[19]

---

[17] A complete model would therefore include an extension of this game, to further stages. In each stage the population that has acted in the previous stage may be hired for some social or business interaction. One who hires a shamed in stage t may herself be shamed and her future opportunity to be hired in stage $t+1$ might be adversely affected. Simplicity commands that we narrow our analysis to a two stage game only.

[18] Some people may enjoy shunning others. Their benefit from avoiding interaction with the liable should therefore be deducted from their costs of searching for the non-liable. The distribution of search costs may therefore account for such benefits as well.

[19] If sellers had full bargaining power then they could fully extract buyers' premium despite excess supply. See Peter Diamond, *A Model of Price Adjustment*, 3 J. ECON. THEORY 156 (1971). Thus, assuming enforcers are shamed-centered, for example, non-shamed sellers would charge a price B whereas shamed would charge B-S. This would be invariant to the rate of shaming. Since bargaining power would usually be distributed between sellers and buyers we would expect our results to carry through. Notice we also assume no entry into and exit out of the sellers' market. If entry and exit were possible, we would expect shamed sellers to clutter in certain markets, since such markets would feature lower expected sanctions, as we explain below.

Second, our early discussion of shaming and private enforcement has distinguished between shamed-centered enforcers and offenders-centered enforcers. In the model, the former would feature a constant benefit function, $B'(u_i)=0$, and a positive shaming cost, $S>0$, whereas the latter would have a down sloping benefit function, $B'(u_i)<0$ and zero shaming costs $S=0$. We analyze these two cases below.

Third, a buyer's search and hiring decisions do not depend on her previous search history. Her decision whether to search for a non-liable seller is stationary. Buyers are therefore partitioned to those who search and those who do not search, given their beliefs about the rate of law compliance among sellers.

B. Analysis

Since buyers can only recognize whether a seller is liable or not there are two possible probabilities of being hired: that of a liable seller, denoted h_L, and that of a non-liable seller, denoted h_N. Therefore, the expected utility of a seller who commits the illegal act, denoted U_i^C, is

$$U_i^C = u_i + v\left(ph_L + (1-p)h_N\right),$$

whereas the expected utility of a seller who does not commit the illegal act, denoted U_i^I, is

$$U_i^I = vh_N.$$

A seller would commit the illegal act if and only if his expected utility would be higher than if he does not commit it,

$$U_i^C > U_i^I. \tag{1}$$

Define a threshold u_T such that all sellers whose utility is higher commit the illegal act, and all sellers whose utility is lower refrain from doing so. Thus,

$$u_T + v\left(ph_L + (1-p)h_N\right) = vh_N \Leftrightarrow$$
$$u_T = vp\left(h_N - h_L\right). \tag{2}$$

Examine now the decision of buyer j whether to search for a non-liable seller, for any given $u_T < \bar{u}$.[20] The buyer's expected utility from hiring a liable seller is

---

[20] Since v<1, in equilibrium it must be that $u_T < 1$.

$$\frac{\int_{u_T}^{1} B(u)du}{1-u_T} - S \equiv B_L(u_T) - S$$

and her expected utility from hiring a non-liable seller is

$$\frac{\int_{0}^{u_T} B(u)du + (1-p)\int_{u_T}^{1} B(u)du}{1-p(1-u_T)} \equiv B_N(u_T,p)$$

Differentiating $B_N(u_T,p)$ with respect to p confirms that it is increasing in p since $B'(u)<0$. As more sellers who have committed the illegal act are caught and shamed, the inference that a non-liable seller has not committed the illegal act is strengthened.

For ease of notation let the expected utility from hiring a seller, without knowing whether he is liable or not, be $\bar{B}$.[21]

The buyer's value from searching until a non-liable seller is found is

$$V = \left(1-p(1-u_T)\right)B_N(u_T,p) + p(1-u_T)V - k_j \Leftrightarrow$$
$$V = B_N(u_T,p) - \frac{k_j}{1-p(1-u_T)}. \tag{3}$$

Since the alternative is to hire a liable seller, the buyer would choose to search for a non-liable seller if and only if

$$V \geq B_L(u_T) - S \Leftrightarrow$$
$$B_N(u_T,p) - \frac{k_j}{1-p(1-u_T)} \geq B_L(u_T) - S \Leftrightarrow \tag{4}$$
$$k_j \leq \bar{B} - B_L(u_T) + S\left(1-p(1-u_T)\right) \equiv k_T(u_T,p) \equiv k_T.$$

Thus, k_T is a threshold cost of search. A buyer would hire a liable seller if and only if her search costs are higher than k_T. Notably, this threshold does not depend on the expected benefit from hiring a non-liable seller. Thus, if p is changed, its only effect on the number of sellers who search depends on the shaming cost, S. We explore this result next:

---

[21] $\bar{B} \equiv pB_L(1-u_T)(u_T) + \left[1-p(1-u_T)\right]B_N(u_T,p)$.

Examine the direct effect of increasing the probability of detection, p, on the two types of private enforcers. Rewriting condition (4) for shamed-centered enforcers we get

$$k_T = S\left(1 - p\left(1 - u_T\right)\right). \tag{4a}$$

Clearly, k_T is decreasing in p. Thus, given a fixed level of seller compliance with the law, u_T, increasing the number of sellers who are shamed reduces the number of shamed-centered enforcers who are willing to incur the costs of search in order to hire a non-liable seller.

Rewriting condition (4) for the offenders-centered enforcers we get

$$k_T = \bar{B} - B_L\left(u_T\right). \tag{4b}$$

Hence, the number of offenders-centered enforcers is *independent* of the number of shamed sellers, given u_T. We summarize these two observations in the following proposition, and then provide some intuition for it.

Proposition 1. *For any fixed level of deterrence, u_T, increasing the probability of detection and shaming, p, would **reduce** the level of enforcement by shamed-centered enforcers, and would have **no effect** on the level of enforcement by offender-centered enforcers.*

To understand Proposition 1 observe that the threshold buyer type, whose search costs are k_T, is indifferent between hiring a liable seller, and searching for a non-liable one. Her payoff from hiring a liable seller in some period t therefore equals her payoff from waiting one period and buying from *any* seller at that next period (since at that period she is also indifferent between searching and not searching if she happens to pick a liable seller). Since an offender-centered buyer incurs no shaming costs, her payoff from buying from a liable seller as well as her expected payoff from hiring any seller to perform the service, irrespective of his liability, are unaffected by the number of shamed. As for a shamed-centered buyer, her payoff from hiring a liable seller does not depend on the number of shamed, yet her payoff from buying next period from any seller decreases since the probability she would have to incur the shaming cost is higher. The offender-centered threshold buyer would therefore remain indifferent between searching and not searching, whereas the shamed-centered threshold buyer

would now prefer not to search. The threshold search costs must therefore decrease in the latter case.

C. Policy Implications

We now turn to analyze sellers' stage 1 strategies, and the policy implications of Proposition 1. As we show, the inverse effect of the number of shamed on the level of enforcement by shame-centered buyers may translate to inverse relations between the rate of shaming and the expected sanction. These may carry over to the three policy instruments that are used to improve deterrence – the *probability* of detection, the *magnitude* of the shaming sanction, and the *accuracy* of shaming. We examine each of these instruments next.

Stage 1 decision of all sellers whether to commit the illegal act is fully characterized by u_T. The probability of a liable seller to be chosen by a buyer is

$$h_L = r\left[1 - G\left(k_T\left(u_T, p\right)\right)\right],$$

and the respective probability of a non-liable seller is

$$h_N = r\left[1 - G\left(k_T\left(u_T, p\right)\right) + \frac{G\left(k_T\left(u_T, p\right)\right)}{\left(1 - p\left(1 - u_T\right)\right)}\right].$$

where the first term in the parenthesis is the probability of being chosen by a buyer whose search costs are too high to search, and the second term is the probability of being hired by a buyer whose search costs are sufficiently low and she therefore searches for a non-liable seller.[22] It is assumed for simplicity that r is sufficiently low, so $h_N > 1$.

Substituting into (2) we get the equilibrium condition (where u_T^* and k_T^* are the equilibrium thresholds),

$$u_T^* = \frac{rvpG\left(k_T^*\left(u_T^*, p\right)\right)}{1 - p\left(1 - u_T\right)} = \frac{rvpG\left(\overline{B} - B_L\left(u_T^*\right) + S\left(1 - p\left(1 - u_T\right)\right)\right)}{1 - p\left(1 - u_T\right)}. \quad (5)$$

On the right hand side of condition (5) we have the *actual* cutoff type, assuming a *belief* that the cutoff would be u_T^*. A Perfect Bayesian (or rational expectation) Equilibrium obtains if and only if the actual cutoff equals the expected cutoff. The proof for the existence of at least one such equilibrium is given in the Appendix (A1).

---

[22] It may be verified that $p(1-F(u_T))h_L + (1 - p(1-F(u_T)))h_N = r$.

Although any mix of shamed-centered and offender-centered motivations can be analyzed in this model, we simplify the following discussion by distinguishing the two extreme cases. Thus, if all buyers are shamed-centered then condition (5) can be rewritten

$$u_T^* = \frac{rvpG\left(S\left(1-p\left(1-u_T\right)\right)\right)}{1-p\left(1-u_T\right)},$$ (5a)

whereas if all buyers are offender-centered then condition (5) would be rewritten as

$$u_T^* = \frac{rvpG\left(\bar{B} - B_L\left(u_T^*\right)\right)}{1-p\left(1-u_T\right)}.$$ (5b)
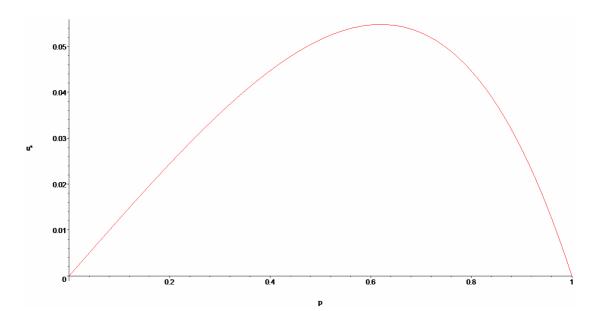
1) Probability of Detection

To analyze the effect of an increase in the probability of detection on the level of crime, examine the equilibrium condition (5). In the Appendix (A2) we show that increasing p affects the level of deterrence in three ways. First, the probability of being detected after committing the illegal act increases. Second, the probability of not being hired, conditional on being detected, decreases. And third, the opportunity to be hired by buyers who only hire non-liable sellers becomes more valuable, since there are less non-liable individuals to choose from. Which of these effects dominates depends on the specific distributions of the benefit from committing the crime and of the costs of search, as well as on the shaming costs incurred by buyers if they hire the shamed.

In particular, if all buyers are offender-centered, the second effect would equal 0, and therefore an increase in p would improve deterrence. The more interesting case is that of shamed-centered enforcers. Here, the second effect may dominate, in which case seller compliance with the law would decline. In fact, the same may be true whenever shaming costs are positive. A direct implication of this result is that maximal compliance with the law may require that the rate of detection and shaming be lower than 1. Proposition 2 summarizes these findings:

Proposition 2. *If the probability of detection, p, is increased then the rate of sellers' compliance with the law would increase if all buyers are offenders-centered (S=0) but it may decrease otherwise. In particular, whenever S>0 maximum seller compliance with the law may require that the probability of detection be lower than 1.*

To take one example where this would indeed be the case, assume $S=1$, $r=0.25$, $v=1$ and $G(k)=\dfrac{k^2}{k^2+1}$.[23] Substituting in (5a) and using a mathematical software to solve we find only one positive solution for each p. Graphing u_T^* as a function of p we get

Clearly, the maximum compliance (highest u_T^*) is obtained in $p<1$.[24]

To sum up, increasing the rate of detection within the frame of a shaming scheme has several results which go in different directions. For our purposes, it is particularly important to notice that increasing the rate of detection decreases the deterrent effects of shaming since it increases the number of shamed individuals in the society and, as was shown earlier, such an increase decreases the expected costs of shaming due to the larger search costs it imposes on law-abiding individuals. Conflicting effects may of course outweigh this effect. The overall effect of increasing the rate of detection depends upon the relative intensity of this effect vis a vis other effects.

2) Magnitude of Sanction

Shaming penalties require the authorities to disseminate information regarding a person who was convicted of a crime. It is often claimed that it is difficult to

[23] See the Appendix **A4** for a graph of the pdf $g(k)$. We take $v=1$ to simplify the example. Clearly, any v sufficiently close to 1 would feature similar results..

[24] The rate of compliance in this example is very low (maximum compliance below 10%). This, however, is a specific feature of the simple distribution function $G(\cdot)$ chosen.

15

disseminate this information in a way that is proportional to the severity of the offence. It is also difficult to measure and control the degree of stigma associated with shaming penalties.[25]

Yet, despite this difficulty there are ways in which the severity of shaming penalties can be controlled or manipulated. The authorities can increase the shaming penalty either by increasing the number of people to whom the relevant information is disseminated (by using more effective means of communication) or by increasing the length of time during which the information is disseminated.[26] In one case, the information concerning any instance of shaming is better disseminated and consequently every law abiding person is aware of more shamed individuals. In the second case, the information concerning every instance of shaming remains longer in the public sphere, and consequently there are more people shamed at any point in time. Our analysis is limited to the first case in which the magnitude of the sanction can be increased, namely by facilitating better dissemination of information.[27]

To examine the sanction's magnitude's effect on deterrence assume that all criminals are announced liable, yet each buyer's (independent) probability of identifying a liable seller is *p*. This case is equivalent to the one analyzed in our model. Similar conclusions therefore follow:

Improving the availability of information concerning liability would increase the number of liable sellers each buyer is aware of. Therefore, the probability of providing the service after committing the illegal act would increase when information is made more publicly available, whenever shaming costs are positive. In equilibrium, however, higher publicity of each seller's liability would also make her more prone to be identified, and would render not committing the illegal act more valuable. Which of these effects would dominate depends on the parameters of the population.

---

[25] See Garvey, *supra* note at 9 at 748 ("[U]nlike fines and imprisonment…[shaming penalties] make the proportionality calculus much more difficult, increasing the risk of disproportionate punishment"); Kahan & Posner *supra* note 2 at 384-85 ; Massaro, *supra* note 4 at 693 (Shaming penalties may hurt *some* offenders too little (because the penalties miss their shame marks), may hurt *some* offenders too much.(because the penalties hit their shame marks), may do nothing *by themselves* to deter norm violations…and may also have unpredictable impact on an offender's subsequent behavior.").

[26] Kahan & Posner, *supra* note 2 at 386.

[27] Analyzing the consequences of the second way of influencing the magnitude of the sanction requires extending the model to a multi-stage setting, which is beyond the scope of this paper. We conjecture, however, that extending 'shaming time' would result in more publicly known liable individuals per-period, and would therefore have similar effects to the ones analyzed in the text.

3) Accuracy of the Judicial Process

The judicial process is not necessarily accurate. Courts may mistakenly acquit guilty individuals (*type I mistake*), or convict innocent individuals (*type II mistake*). An individual's decision whether to commit the illegal act would depend on judicial accuracy – the difference in the expected sanction if he commits the illegal act and if he does not.

In the context of traditional sanctions, an increase in judicial accuracy would either increase the probability that an offender will be convicted, or increase the probability that an innocent person be acquitted, or both. Consequently, an increase in judicial accuracy would increase the gap in the expected costs resulting from the imposition of criminal sanctions between offenders and non-offenders. By convicting more criminals (rather than exculpating them by mistake) and by acquitting more innocent individuals (rather than convicting them by mistake), the judicial system would increase the expected costs of sanctions for criminals relative to those imposed on non-criminals. Hence, such an increase in the accuracy of the judicial process would be conducive to deterrence.

Moreover, fixing this difference (and consequently, the level of accuracy), deterrence would not depend on the absolute rate of false convictions or false acquittals. For example, when the probability of convicting a guilty individual is 0.8 and the respective probability for an innocent individual is 0.3, the level of deterrence would the same as when the respective probabilities are 0.7 and 0.2, as the difference between the two probabilities is kept constant (0.5). That the first case has more type II mistakes and less type I mistakes would not affect deterrence.[28] We now show that these results may not hold in the context of shaming sanctions.

Since the model assumed that not all sellers who committed the illegal act are found liable, it implicitly allowed for one type of inaccuracy, namely mistaken acquittal (type I mistake). Yet, the other type of inaccuracy, mistaken conviction (type II mistake), was not taken into account. In the Appendix (A3) we extend the model for this type of mistake by letting the probability of finding an innocent individual liable $q>0$.

---

[28] See Louis Kaplow & Steven Shavell, *Accuracy in the Determination of Liability*, 37 J. L. & ECON. 1 (1994); Louis Kaplow, *The Value of Accuracy in Adjudication: An Economic Analysis*, 23 J. LEGAL STUD. 307 (1994).

We show that the relationship between accuracy and deterrence changes when shame sanctions are used. Unlike traditional sanctions deterrence resulting from shame penalties is not fully specified by the level of accuracy (p-q). Fixing the level of accuracy, and increasing the probability of finding both an innocent and a liable seller liable decreases k_T^*, thus decreasing the ratio of buyers who search for a non-liable seller. Yet, at the same time it increases non-liable sellers' expected probability of being hired by buyers who search, conditional on being chosen by such a buyer. Whereas the first effect undercuts deterrence, the latter improves it. Thus, for any level of accuracy, the decision whether to convict or acquit more criminals should depend on which of these two effects dominates. The level of deterrence would therefore not be invariable to the rate of type I and type II mistakes, even if the difference between the probability of guilty and innocent sellers to be found liable is kept constant.[29]

4. From Theory to Practice: Existing Legislation and the Search Model

The discussion above specified several effects that an increase in the probability of shaming would inevitably have. First, an increase in the rate of detection would have a direct effect, increasing the expected sanction for those who commit the illegal act by increasing the probability they would be detected and shamed. Second, as a result of such an increase, the rate of enforcement by shamed-centered enforcers would decline.[30] This second effect is absent if and only if all enforcers are offenders-centered. Thus, the sanction a shamed individual expects if he were to only interact with offenders-centered enforcers would increase due to the direct effect of the higher rate of detection. Yet, if the shamed individual expects interaction with shamed-centered enforcers as well, the second (and opposite) effect would kick in. Which of these effects would dominate is difficult to predict absent a concrete specification of the way the offender and enforcer populations are distributed (according to their utility from committing the offence and their search costs,

---

[29] For example, changing the burden of proof may affect deterrence even if the consequent change in the rate of convicted individuals is the same among criminals as among innocent (that is, even if the level of accuracy is kept unchanged).

[30] As we show in our model there is actually a third effect, which is more refined: as the rate of shaming increases, the probability of social or business interaction for those who were not shamed increases, thus rendering committing the offense relatively more costly. See section 3.C.1) above.

respectively) and the ratio of offenders-centered and shamed-centered enforcers. [31] It is important therefore to establish that

To examine the efficacy of current shaming legislation schemes, we now focus on the role shamed-centered individuals play in them. We suggest that shaming schemes rely significantly on the cooperation of shamed-centered individuals, thus rendering them problematic, following above analysis. We then discuss one of the most compelling arguments favoring the use of shaming penalties, namely their expressive value. Since the expressive value of shaming penalties relies, at least partly, on cooperation of shamed-centered enforcers, and since such cooperation may decline with the rate of shaming, a higher rate of shaming might, paradoxically, adversely affect the expressive value of the law.

A glance at the offences which are currently punished by shaming penalties demonstrates that many of these sanctions rely primarily on the cooperation of shamed-centered individuals. Whitman classifies offences that are currently punished by using shaming penalties into three types: 1) sexual and moral offences; 2) commercial offences; and 3) first and minor offences.[32]

Some sexual offenders would be subjected to sanctions by offenders-centered shamers. Thus, for instance, the declared purpose of Megan laws is to provide information to parents who can better protect their children on the basis of the information provided in accordance with the Megan laws.[33] The unwillingness of parents to associate with sexual offenders is grounded in their conviction that sexual offenders, in particular child molesters, threaten the safety of their children.

---

[31] To better understand our claim, it must be distinguished from another related but distinct argument - that there is an inverse relationship between beliefs concerning the rate of crime in the society and the effectiveness of shaming. Dan M. Kahan, *Social Influence, Social Meaning and Deterrence* 83 VA. L. REV. 349, 357 (1997) ("The more prevalent criminal activity is in a particular community the less likely someone is to be condemned for it by either those with criminal records or those without"). For the claim that the rate of crime can influence the size of the social sanction, see Oren Bar-Gill, Alon Harel, *Crime Rates and Expected Sanctions: The Economics of Deterrence Revisited* 30 J. LEGAL STUDIES 485 (2001). In contrast, our argument is not founded on the premise that the effectiveness of shaming penalties depend on the beliefs concerning the *rate of crime*; instead, it is founded on the claim that the effectiveness of shaming penalties depend on the *rate of the infliction of shaming penalties* (*i.e., on the number of shamed individuals*) in the society. The more people are shamed the less effective shaming penalties are. Our model is based on the rate of **shaming** and not on the rate of crime.

[32] See James Q. Whitman, *What is Wrong with Inflicting Shame Sanctions?* 107 YALE L.J. 1055, 1064 (1998).

[33] It is on the basis of the claim that these laws are intended to warn the public and not to impose additional punishment that these laws have been upheld against constitutional challenges. See, e.g., Lieb, R., Quinsey, V., and Berliner, L., *"Sexual Predators and Social Policy,"* in M. Tonry (Ed.), CRIME AND JUSTICE 43, 78-79 (1998)

Increasing the rate of shaming provides valuable information to these parents and the effects of increase in the search costs is outweighed by the superior quality of the information. Yet other types of sexual offences do not seem to raise this type of fear. Thus, for instance, the efficacy of the practice of publicizing the names of patrons using the services of prostitutes seems to rely on the cooperation of shamed-centered enforcers.[34] The goods or services provided by johns are not likely to be inferior to those provided by others and the unwillingness to associate with johns seems to fit the patterns characterizing shamed-centered offenders.

Shaming penalties imposed on commercial offences also fall in both categories. Some commercial offences such as fraud can undermine one's reputation. Merchants and participants in the world of commerce inevitably fear loss of reputation – reputation that is directly related to the quality and reliability of goods and services expected from them by potential customers. Yet shaming penalties have also been applied in the context of tax evasions.[35] In these cases, shaming is more likely to rely on the cooperation of shamed-centered shamers. The quality of goods or services provided by tax evaders is not lower than that provided by others. Increasing the rate of shaming increases the search costs without providing tangible benefits to law-abiding individuals.

The last category – the category of minor offenders -- such as shoplifters and drunk drivers seems also to rely on the cooperation of shamed-centered shamers. When officials inflict a shaming sanction on a person in this category – circulating a shoplifter photograph or requiring a drunk driver to wear a pink bracelet reading D.U.I CONVICT, they aim at inducing the person's friends and acquaintances to shun him not because that behavior is indicative that the person is disloyal to his friends or poses risk to their well-being. Associating with these individuals indicate lack of respect toward the relevant legal norms. An increase in the rate of shaming of offenders belonging to this category may therefore have counterproductive effects.

Moreover, shaming penalties are often praised not merely because of their efficacy in deterring criminals. It is often asserted that shaming penalties are superior in that they are expressive of the community's moral disapproval. Under this view,

---

[34] For a survey of the practice of publishing names and pictures of prostitutes' patrons, see: *Sex in the Sunlight: The Effectiveness, Efficiency, Constitutionality, and Advisability of Publishing Names and Pictures of Prostitutes' Patrons* 49 VAND. L. REV. 1525 (1996)

[35] See Michael S. Kirsch, *Alternative Sanctions and the Federal Tax Law: Symbols, Shaming and the Social Norm Management as a Substitute for Effective Tax Policy* 89 IOWA L. REV. 863 (2004)

shaming penalties are particularly appropriate to convey repugnance towards the criminal act. They "denounce the wrongdoer and his conduct as contrary to shared moral norms; and they ritualistically separate the wrongdoer from those who subscribe to such norms".[36]

Yet, since increasing the rate of shaming increases the proportion of shamed-centered enforcers who fail to cooperate with the shaming scheme, the normative message conveyed may be diluted. Although the law denounces the offender, less members of the community follow suit, thus questioning whether the violated norm is indeed shared by the community. A dissonance is produced, between the law's disapproval of the illegal act, and the willingness of individuals to overlook it. Increasing the rate of shaming may therefore fail not only in substituting for traditional sanctions' deterrent functions, but also in reinforcing community's cooperation with the law. Thus, an increased rate of shaming may paradoxically undermine the law's expressive value.

5. Conclusion

In a traditional sentencing scheme based on imprisonment and fines, an increase in the rate of detection of criminals inevitably leads to an increase in the expected costs of the sanction. Similarly, increasing the sentence inevitably leads to an increase in the deterrent effects of the sanctions. These two observations provide the basis for the economic analysis of crime and law enforcement.[37] In addition, both common sense and law and economics scholars suggest that increasing the accuracy of the judicial system leads to an increase in the expected costs of committing a crime and consequently reduces the rate of crime. This paper showed that at least in the context of penalties which derive most of their power from publicizing the criminals' conviction and from the resulting isolation and alienation of criminals, these traditional claims may be false. Increasing the rate of detection, the magnitude of the sanction and the accuracy of the judicial process may decrease rather than increase deterrence. Shaming penalties are perhaps the prime example of penalties which acquire their deterrent effects from the publicity resulting from their imposition.

---

[36] See Kahan, *supra* note 2, 636.
[37] Gary Becker, *Crime and Punishment: an Economic Approach*, 76 JOURNAL OF PUBLIC EECONOMICS 169 (1968)

Hence, these perverse effects are most likely to occur in the context of shaming penalties.[38]

While this paper does not directly take sides in the debate between advocates and opponents of shaming penalties, its conclusions are ones that suggest that the use of shaming penalties is restricted. If the deterrence resulting from the imposition of shaming penalties is insufficient, it ought to be complemented by greater use of regular sanctions rather than by greater use of shaming penalties. Shaming penalties can be 'self destructive' as an extensive use of them may erode their effectiveness.

---

[38] Search costs provide but one mechanism to explain the inverse relations between the rate of shaming and the deterrent effects of shaming penalties. There are at least two other mechanisms that may have similar effects. One is a *bounded information model* grounded in the cognitive limitations of "shamers"; the other is a *group formation model,* that features a possible constraint on the number of shamed individuals who are socially sanctioned. Under the Bounded Information Model, the more people are shamed, the lesser the social isolation of the shamed is since the ability of law abiding individuals to identify the shamed and isolate them decreases. Under the Group Formation Model, The more people shamed, the larger the ability of the shamed to form alternative communities and consequently, the lesser the expected costs of being shamed. Moreover the more people shamed, the lesser the ability of law-abiding individuals to form law-abiding communities. Further research is called for to examine the implications of these two models.

Appendix

A1.

Define the (continuous) function $H(u_T)$ as follows

$$H\left(u_T\right) = \frac{rvpG\left(k_T\left(u_T, p\right)\right)}{\left(1 - p\left(1 - u_T\right)\right)},$$

where $k_T(u_T, p)$ is defined according to (4). Clearly, $H(u_T) > u_T$ when $u_T = 0$. Also, since $v < 1$ it must be that $\lim_{u^* \to 1} H\left(u_T\right) < u_T$. By the intermediate value theorem there must be $u_T = u_T^*$ such that $H(u_T^*) = u_T^*$.

A2.

Define a *locally stable equilibrium* to be a Perfect Bayesian Equilibrium[39] of the game (u^* satisfying condition (5)), for which the adjustment process

$$u_{t+1}^* = \frac{rvpG\left(k_T^*\left(u_t^*, p\right)\right)}{\left(1 - p\left(1 - u_T\right)\right)}, \; t = 0, 1, 2, \dots \text{ converges to } u_T^* \text{ for } u_0^* \text{ sufficiently close to } u_T^*.[40]$$

An equilibrium is locally stable only if $\dfrac{\partial H\left(u_T^*\right)}{\partial u_T^*} < 1$. By the argument in A1 there must be at least one locally stable equilibrium in this game.

Totally differentiating condition (5) with respect to p we get

$$\frac{du_T^*}{dp} = \frac{\partial H\left(u_T^*\right)}{\partial p} \bigg/ \left(1 - \frac{\partial H\left(u_T^*\right)}{\partial u}\right).$$

Since in a locally stable equilibrium $\dfrac{\partial H\left(u_T^*\right)}{\partial u_T^*} < 1$, $\dfrac{du_T^*}{dp}$ has the same sign as $\dfrac{\partial H\left(u_T^*\right)}{\partial p}$.

Differentiating (5a) we get

$$\frac{\partial H\left(u_T^*\right)}{\partial p} = rv \left\{ \frac{G\left(S\left(1 - p\left(1 - u_T\right)\right)\right)}{\left[1 - p\left(1 - u_T\right)\right]^2} - \frac{pS\left(1 - u_T\right)G'\left(S\left(1 - p\left(1 - u_T\right)\right)\right)}{1 - p\left(1 - u_T\right)} \right\},$$

which may be negative for shamed-centered enforcers if S>0 (see example in text). In particular, this derivative may be negative for $p = 1$.

---

[39] For a definition of a Perfect Bayesian Equilibrium see Osborne & Rubinstein (1995), p. 231.

[40] For a similar definition see Rasmusen, *supra* note 5; See also Stephen Coate and Glen Loury, *Will Affirmative Action Policies Eliminate Negative Stereotypes*, 83 **AMERICAN ECONOMIC REVIEW** 1220, 1226 (1993).

A3.

In this part of the Appendix we modify our analysis for the case where innocent individuals may be found liable by mistake. We assume all sellers are audited, and let p and q denote the probability of finding a guilty and an innocent sellers liable, respectively. Thus, the expected utility of a guilty seller is

$$U_i^C = u_i + v\left(ph_L + (1-p)h_N\right),$$

whereas the expected utility of an innocent seller is

$$U_i^I = v\left(qh_L + (1-q)h_N\right).$$

Thus,

$$u_T + v\left(ph_L + (1-p)h_N\right) = v\left(qh_L + (1-q)h_N\right) \Leftrightarrow$$

$$u_T = v(p-q)\left(h_N - h_L\right).$$

(A1)

If a buyer picks a liable seller his expected utility from hiring her to provide the service is

$$\frac{q\int_0^{u_T} B(u)du + p\int_{u_T}^1 B(u)du}{qu_T + p(1-u_T)} - S \equiv B_L\left(u_T, p, q\right) \equiv B_L,$$

and her expected utility from hiring a non-liable seller is

$$\frac{(1-q)\int_0^{u_T} B(u)du + (1-p)\int_{u_T}^1 B(u)du}{(1-q)u_T + (1-p)(1-u_T)} \equiv B_N\left(u_T, p, q\right) \equiv B_N$$

The buyer's value from searching if she picked a liable seller is

(A2)

$$V = (1-q)\int_0^{u_T} B(u)du + (1-p)\int_{u_T}^{\bar{u}} B(u)du + \left[qu_T + p(1-u_T)\right]V - k_j \Leftrightarrow$$

$$V = B_N - \frac{k_j}{(1-q)u_T + (1-p)(1-u_T)}$$

Since the alternative is to hire the liable individual, the employer would choose to search for a non-liable individual if and only if

$$V \geq B_L - S \Leftrightarrow$$

$$B_N - \frac{k_j}{1 - qu_T - p(1 - u_T)} \geq B_L - S \Leftrightarrow$$

$$k_j \leq \bar{B} - B_L + S(1 + (p - q)u_T - p) \equiv k_T(u_T, p, q) \equiv k_T.$$

Whereas h_L is defined similarly to the case where no innocent individuals are found liable,[41] h_N is now given by
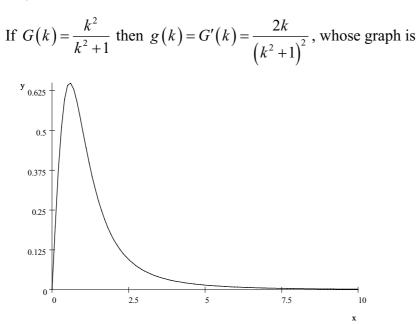
$$h_N = r\left[\left(1 - G(k_T(u_T, p))\right) + \frac{G(k_T(u_T, p))}{\left(1 - qF(u_T) - p(1 - F(u_T))\right)}\right]$$

The equilibrium condition is therefore

$$u_T^* = \frac{rv(p - q)G(k_T^*)}{1 - p(1 - u_T^*) - qu_T^*} = \frac{rv(p - q)\left(G\left(\bar{B} - B_L + S(1 - qu_T^* - p(1 - u_T^*))\right)\right)}{1 - p(1 - u_T^*) - qu_T^*} =$$

$$= \frac{rv(p - q)\left(G\left(\bar{B} - B_L + S(1 + (p - q)u_T^* - p)\right)\right)}{1 + (p - q)u_T^* - p}.$$

---

[41] See *supra*, p. 13.

A4.

If $G(k) = \dfrac{k^2}{k^2+1}$ then $g(k) = G'(k) = \dfrac{2k}{\left(k^2+1\right)^2}$, whose graph is

26