

On the Interaction between Legal and Reputational Sanctions

Edward M. Iacobucci

Faculty of Law, University of Toronto

January 2014

[edward.iacobucci@utoronto.ca](mailto:edward.iacobucci@utoronto.ca)

Thanks to participants at the Berkeley Law and Economics Workshop, American Law and Economics Association Annual Meeting, the Canadian Law and Economics Association Annual Meeting, Siena-Toronto-Tel Aviv workshop for helpful comments. Thanks also to Omri Ben-Shahar, Ralph Winter and an anonymous referee for their comments.

**ABSTRACT**

This article shows that reputational sanctions are not, as the literature implicitly assumes, independent of legal penalties. Rather, observers will assess a wrongdoer's type in light of conduct and the expected legal sanctions for that conduct. For example, if a person commits a wrong even in the face of onerous legal penalties, observers will tend to draw a more negative inference about the wrongdoer's propensity for wrongdoing than if legal penalties were light; on the other hand, onerous legal penalties may deter wrongdoing even for actors with a relatively high propensity for wrongdoing, which also affects reputational penalties. In addition, the article demonstrates that the literature's focus on deterrence in considering the relationship between formal and informal penalties is too narrow. It may be socially preferable in some circumstances to adjust legal penalties to allow actors to reveal their types than to adjust legal sanctions to promote optimal deterrence.

## I. INTRODUCTION

There is a large literature on the interaction between formal, legal sanctions, and informal, non-legal sanctions. Much of the literature analyzes the substitutability of legal and non-legal sanctions, pointing out that various non-legal sanctions such as stigma, or loss of standing in a community, may deter undesirable behaviour just as, or more, effectively than formal legal sanctions (see, e.g., Macauley (1963), Ellickson (1991), Bernstein (1992)). Other aspects of the literature focus on the potential complementarity of informal and formal sanctions, noting that legal penalties may influence the existence and impact of informal sanctions (see, e.g., Posner (2000), Cooter (2000), Teichman (2005), Baniak and Grajzl (2013)). This article engages a subset of these broader debates by asking the question, to what extent should the size of a legal penalty be adjusted in light of non-legal sanctions? There are obviously significant informational hurdles (e.g., what will the cost of reputational sanctions be?) that would make such adjustments very difficult in practice (Cooter and Porat (2001); this article shows that the problem is more complicated as a conceptual matter than has been appreciated.

The distinction between two kinds of informal penalties is not always drawn clearly in the literature, but it is crucial for analyzing the reputational effects of law. On the one hand are penalties that arise from the conscious, intentional punishment of a wrongdoer by members of a community. The other kind of informal sanction results from a loss to reputation. In some ways all informal sanctions can be conceived as reputational in nature. Ostracism, for example, results from a stigma attaching to a wrongdoer that hurts that person's reputation and standing in the community. What I will call a reputational sanction arises because of the observer's pure self-interest, not a self-interest derived from wanting to adhere to a social norm, or to avoid punishment from others. The reputational sanction arises because observers have changed their

views about the benefits of dealing with a wrongdoer that has revealed by its wrong its type as one that is unattractive to trading partners (with “trading” conceived broadly).

Commentary has explored the complementarity between informal and formal sanctions largely in the context of stigma. Formal litigation may produce information that is relevant to informal sanctions (see, e.g., Posner (2000), Baker and Malani (2011), Shapira (2013)). The law may also have an expressive function, which interacts with informal sanctions and stigma (see, e.g. Lessig (1995), Cooter (2000), Cooter (2000a), Cooter (2000b), Scott (2000), Teichman (2005), Posner (2000), McAdams (2000), Kahan and Posner (1999), Kahan (1997), Baniak and Grajzl (2013)). While both stigma and reputational penalties are important in different circumstances, I focus in this paper on purely reputational penalties. Such reputational penalties are particularly apt where there are diffuse, atomistic participants in a market. Collective action problems affect stigma, since punishment costs the punisher, but do not affect purely reputational sanctions since observers punish out of their own narrow self-interest. Self-interested punishment, which is not undermined by collective action problems, is the more plausible kind of sanction in commercial markets with many participants.

A number of empirical studies have shown that there are significant informal sanctions that result from a wide range of wrongs (see, e.g., Peltzman (1981), Mitchell and Maloney (1989), Karpoff et al. (2008), and Armour et al. (2010)). Several commentators conclude that the informal losses are only significant where the victim from the misconduct is a specific counterparty with the wrongdoer, and not where the wrong is to third parties or the public at large (see, e.g., Karpoff and Lott (1993), Alexander (1999), Karpoff et al (2005), Murphy et al. (2009), and Armour et al. (2010)). This provides support for the conclusion that wrongs signal something

undesirable about the wrongdoer's type, which costs the wrongdoer trading opportunities going forward.

Taking as given the proposition that informal sanctions follow from the imposition of formal penalties, several commentators have suggested that the law should account for informal sanctions when imposing formal sanctions. Cooter and Porat (2001), for example, argue that optimal legal penalties should account for the effects of informal penalties. Karpoff and Lott (1993) and Polinsky and Shavell (2011) suggest that legal penalties may not be appropriate at all in some settings because of the size and significance of informal sanctions.<sup>1</sup>

The literature on this question has overlooked an important interaction between formal and informal penalties: the reputational sanction that the wrongdoer bears from certain conduct will generally *depend* on the size of the legal penalty associated with that conduct. The overlooked intuition is basic. For there to be reputational implications of certain conduct, it must be that the potential wrongdoer's choice to engage in misconduct will be influenced by its type. But legal sanctions will also affect choices. If a wrongdoer is willing to act in a certain way *despite* a significant expected legal sanction for such behaviour, then the wrongdoer is of a type that is especially willing to engage in such behaviour. The intuition might be that the higher the legal sanction, the more severe the reputational sanction, but this is misleading. The reputational sanction results from the comparison of a wrongdoer with a non-wrongdoer. A more severe legal sanction affects both comparators: wrongdoers will be perceived on average as especially likely to commit the wrong as the legal penalty increases; but as more types are deterred by the increasing sanction, the pool of non-wrongdoers deteriorates and non-wrongdoers will also be

---

<sup>1</sup> Ganunza et al. (2011) challenge this idea by noting that informal sanctions may be more costly than legal sanctions. Their analysis depends on the stigma conception of informal harms that involves punishers potentially suffering costs, rather than the pure reputational conception that I rely upon (and believe is more plausible in commercial settings with many different actors).

perceived to be more likely to commit the wrong on average as the penalty increases. Since increasing the legal sanction affects both sides of the comparison between wrongdoers and non-wrongdoers, the effect on the reputational sanction is ambiguous.

The article proceeds as follows. Part II develops the article's main argument about the effect of formal sanctions on reputational sanctions. Part III reviews the policy implications of the argument, showing that, in addition to deterring wrongs, the law may make a valuable social contribution by screening types. For this reason, there is an additional consideration about the value of increasing the probability of detecting wrongdoing that is relevant to the literature on law enforcement (Becker 1968); Part IV makes this point. The analysis shows that accounting for reputational sanctions in setting legal sanctions is more complex than the literature has recognized. Part V concludes.

## II. REPUTATIONAL PENALTIES DEPEND ON LEGAL PENALTIES

The literature's analysis of the interaction between law and reputation has failed to appreciate an important effect of the legal sanction on the reputational sanction. This section shows that the reputational losses from particular misconduct will generally be contingent on the penalty, not because of group norm enforcement, but rather because the signal about type sent by misconduct will vary with the legal penalty that the misconduct attracts. Reputational and legal penalties interact in a way that the literature has not recognized. This section shows that increasing the legal sanction may increase or decrease the reputational sanction.

To reiterate, I focus on the reputational effect that arises where a legal penalty allows others to learn of the wrongdoer's propensity for certain conduct, which affects their willingness to deal with the wrongdoer. There are any number of possible propensities that wrongdoing could reveal. It could reveal that the wrongdoer accounts little for harmful effects on others; this

would make others less willing to deal with the wrongdoer. A wrong, and a willingness to sacrifice reputation as a consequence, could also reveal a high discount rate such that the wrongdoer prefers immediate gain to future gains.<sup>2</sup> It could also be that there is a cost of not causing harm to others in a particular circumstance, and wrongdoers generally have higher costs of taking care of others. These characteristics could apply to an individual, or to a corporation. The corporation would have these tendencies perhaps because of characteristics of the business, such as the firm's cost of capital and consequent discount rate, its access to technology, or because of the characteristics of its managers as individuals and their freedom to act to at least some extent in their own interests because of principal-agent problems (see, e.g., Iacobucci (2006)).<sup>3</sup>

In the simple model that I rely on I assume that there is a class of prospective wrongdoers that vary in their costs of complying with a particular law. There is a legal sanction associated with a failure to comply but wrongdoers will only comply if the benefits of doing so outweigh the costs. Given my reliance on a reputational sanction that results from other parties' reluctance to deal with the wrongdoer, I implicitly consider laws that concern misconduct towards a party with which the wrongdoer deals directly. Think, for example, of laws against misrepresentation in securities markets, rather than laws seeking to limit money laundering (see Armour et al. 2010).

The timing of the model is as follows. In the first period, the prospective wrongdoer (the "firm") must choose whether to comply with a law or not. If the firm fails to comply with the

---

<sup>2</sup> Posner (2000) focuses on this propensity.

<sup>3</sup> Of course, if corporate attributes that give rise to a potential reputation can be costlessly and instantly changed, corporate reputations might not necessarily form: observers would not necessarily trust what they observe anticipating that what they observe can easily be changed. The evidence, however, is that there are reputational effects on corporations from wrongdoing, suggesting that there is reason to view reputationally relevant characteristics as sticky even in the corporate context.

law, it faces an expected legal sanction,  $pF$ , where  $p$  is the probability of detection, and  $F$  is the fine. If the firm complies with the law, it incurs a cost of compliance,  $C$ . For simplicity, assume that there are two types of firms: “good” firms with a relatively low cost of compliance,  $C^G$ , and “bad” firms with a relatively high cost of compliance,  $C^B$ , where  $C^G < C^B$ . Firms are assumed to make a non-negative profit in the first period regardless of their choice to comply. I do not model these positive first period profits explicitly but instead focus only on what in the first period matters to the analysis: the effect on first-period profits from either incurring the costs of compliance, or the expected fine from non-compliance.

After the first period, the firms will be affected by their earlier conduct and consequential reputation. For simplicity, I will rely on a single-shot second period game. This is a natural simplifying assumption given that type is given and immutable.<sup>4</sup> In the second period, firms earn a profit from their activities that is contingent on their reputations from the first period. The second period activity need not be identical to that in the first.<sup>5</sup> What is necessary is that the traits that underlie actions in the first period are relevant to observers in the second. Reputational effects are especially important where the activity in the second period is not directly regulated by law. For example, the first period could involve selling a product under an only imperfectly enforced contractual obligation to provide a certain quality, while the second period involves the

---

<sup>4</sup> The second period could of course itself be conceived of as a repeated game. For example, actors may require a certain minimum reputation at the beginning of a repeated game to provide a sufficient bond against committing wrongs in the future that cooperation in the game is possible (see, e.g., Klein and Leffler 1981). For my purposes, a single-shot second period is sufficient, but in more complex models (e.g., one that accommodates a potential change over time in the actor’s type), this assumption would be inappropriate.

<sup>5</sup> Where the activity is exactly the same in each period, reputational effects may be trivial. For example, in a two-period interaction (or any discrete number of periods), backwards induction implies that the firm will never comply where there is cost of compliance unless the law deters wrongdoing: it will not comply in the last, since there is no future to attend to; which means that it will not comply in the second-last period; and so on. There is no reputational incentive. Moreover, in an infinitely repeated game, there may be signaling and reputational effects, but they will overlap with deterrence considerations and there will not be any social benefit of signaling since one period is identical to the next; firms will either be deterred or not from wrongdoing. I discuss the social benefits of signaling at greater length in Part III.



sale of a product of a quality that is not regulated at all under contract, but involves an extra-contractual quality commitment. Both periods are affected by the firm's cost of providing quality, the lower the cost of providing quality, the better the firm, but legal enforcement is only available to deter misconduct in the first period. I assume for the moment that legal enforcement is not necessary for observers to know whether the firm has complied in the first period. For example, conduct is always observable, but it is only verifiable with probability  $p$  (which is assumed to be known to observers) and thus only probabilistically subject to punishment.

If the firm has a reputation for being good, it earns high second period profits of  $\Pi^G$  (discounted to the first period, as are the other possible profit outcomes), if it has a reputation for being bad, it earns low second period profits  $\Pi^B$ , and if observers cannot tell type from actions in the first period, it earns  $\Pi^P$  in the second period (where  $P$  signifies a pooling equilibrium in which types are indistinguishable). The assumption is that,  $\Pi^G \geq \Pi^P \geq \Pi^B$ .<sup>6</sup>

The firm faces a straightforward decision in the first period: comply where the costs of doing so are smaller than the benefits. The firm of type  $i$  will comply if,  $\Pi^c - C^i > \Pi^{dnc} - pF$ , where  $\Pi^c$  indicates profits in the second period having complied in the first, and  $\Pi^{dnc}$  indicates profits in the second period having not complied in the first.

$\Pi^{dnc}$  and  $\Pi^c$  depend on the nature of the equilibrium. Separating or pooling equilibria may emerge. In a separating equilibrium, good firms and only good firms comply in the first period, which in turn implies that observers will infer that a firm is good if they comply, while

---

<sup>6</sup> It is entirely possible that profit could be a function not only of reputation, but also of actual type. Conditional on a given reputation, high compliance cost types could have higher or lower profits than low compliance cost types. For example, if good firms find it lower cost to be honest, they are more likely to incur costs to provide high quality in the second period, which could imply lower second period profits for good firms conditional on a good reputation. On the other hand, if good firms find it cheaper to provide high quality, it is possible that their profits are higher than bad firms, who may have high costs even to provide a basic quality in the second period conditional on a good reputation. Having profit vary with type complicates the exposition without changing the relevant arguments so I make the simplifying assumption that second period profits are conditional only on reputation.

bad firms do not comply and observers learn their type as a consequence. I assume that all firms earn a positive profit from their activities, some (unmodelled) profits in the first and non-negative profits in the second, such that the participation constraint is always satisfied.

The necessary incentive compatibility conditions for a separating equilibrium are as follows. First, the good firms must find it profitable to comply, in part because of the benefits of a good reputation. That is,  $\Pi^G - C^G > \Pi^B - pF$ . Second, the bad firms must not find it profitable to comply and thus imitate a good firm. That is,  $\Pi^G - C^B < \Pi^B - pF$ . If these conditions are met, and if observers expect good firms to comply and bad firms not to comply, then there is a reputational penalty from failing to comply. Firms that do not comply lose  $\Pi^G - \Pi^B$  in profits in the second period, the reputational sanction, as well as a legal sanction in the form of an expected fine of  $pF$ . Bad firms are willing to suffer the reputational loss and the fines, however, because the reputational loss and the fines are low relative to their high costs of compliance. Good firms prefer to avoid the fines and the reputational sanctions because they are high relative to their low costs of compliance.

A pooling equilibrium could involve either both good and bad firms complying, or both good and bad firms not complying. The compliance equilibrium results where  $\Pi^P - C^G > \Pi^B - pF$  and  $\Pi^P - C^B > \Pi^B - pF$ ; <sup>7</sup> of course the second condition is what matters since  $C^G < C^B$ . In this equilibrium, the profits from maintaining the pooled reputation relative to a bad reputation and the expected fines are sufficiently high, and the costs of compliance sufficiently low, to induce even bad firms to comply such that there is no observable distinction between firms.

---

<sup>7</sup> To determine the reputational sanctions and thus conditions for the pooling equilibrium precisely, one would need to specify observers' beliefs about firm type for strategies that are off the equilibrium path. Since there are only two types, and bad firms find it less profitable to comply, I assume that the observers would assume a bad firm if they observe non-compliance in a compliance pooling equilibrium; below I assume that observers would assume a good firm if they observe compliance in a no-compliance pooling equilibrium.

The no-compliance equilibrium arises where both firms do not comply, which results if  $\Pi^G - C^G < \Pi^P - pF$  and  $\Pi^G - C^B < \Pi^P - pF$ ; the first condition is what matters since  $C^G < C^B$ . In this equilibrium, the losses from forgoing a good reputation relative to a pooled reputation and the expected fines are sufficiently low, and the costs of compliance sufficiently high, to induce even good firms not to comply and there is no observable distinction between firms.

The reputational sanction for particular conduct will depend on the nature of the equilibrium.<sup>8</sup> In all cases the firm must consider the costs (forgone profits) from a lesser reputation that results from not complying, but the size of those forgone profits depends on the nature of the equilibrium. In a separating equilibrium, the cost of not complying is the full difference between the second-period profits of known good firm and the profits of a known bad firm. The reputational sanction is  $\Pi^G - \Pi^B$ . In a pooling equilibrium, in contrast, the cost of not complying when other firms comply is, at most,<sup>9</sup> the difference between a firm's profits when its type is unknown and a known bad firm's profits,  $\Pi^P - \Pi^B$ . Alternatively, in a pooling equilibrium in which all do not comply, the cost of not complying is at most the difference between a known good firm's profits and a firm's profits when its type is unknown,  $\Pi^G - \Pi^P$ .

It is clear that the reputational sanction from not complying is larger in a separating equilibrium than in either of the pooling equilibria. This is because observers infer type perfectly and profits in the second period would drop all the way from those of a good firm to those of a bad firm following non-compliance. In a pooling equilibrium, in contrast, the difference in

---

<sup>8</sup> Note that there are multiple equilibria in some cases. There is an interaction between observers' expectations and firm behaviour. Over some range, if observers expect a separating equilibrium, then good firms would comply in order to get the benefits of a good reputation; if observers expect a pooling equilibrium, in contrast, then good firms may not comply as the reputational benefits are too small to induce compliance.

<sup>9</sup> As noted above, the reputational sanction depends on off-equilibrium beliefs of observers about type. At best, observers would assume a good firm if they were to observe compliance in a non-compliance pooling equilibrium

profits from non-compliance are only between the profits of either a good or bad type and the profits of a firm of average type.

To show that the size of the legal fine and the reputational penalty are related, then, it remains only to show that the size of the fine affects the nature of the signalling equilibrium. It is straightforward to show that the size of the fine indeed does affect the signalling equilibrium, and thus affects the size of the reputational sanction. Suppose as an initial matter that there is a no-compliance pooling equilibrium, in part because there is no fine for committing the wrong; that is  $F=0$ , and even a good type finds the reputational sanction too small to comply:  $\Pi^G - C^G < \Pi^P$ . In this pooling case the reputational sanction from failing to comply is no larger than  $\Pi^G - \Pi^P$ .

Now consider a higher fine,  $\bar{F} > 0$ . If  $\bar{F}$  is high enough, and if the probability of detection is strictly positive, the pooling equilibrium will be replaced by a separating equilibrium. That is,  $\Pi^G - C^G > \Pi^B - p\bar{F}$  and  $\Pi^G - C^B < \Pi^B - p\bar{F}$ . Increasing the fine from zero to  $\bar{F}$  increases the reputational sanction from  $\Pi^G - \Pi^P$  to  $\Pi^G - \Pi^B$ . The intuition is that when fines are very low, nobody finds it profitable to comply, which means that there is less of a penalty from misconduct since everybody does it. In contrast, when fines increase, good firms eventually comply, thus revealing type and increasing the reputational sanction for non-compliant, bad firms.

This analysis is sufficient to demonstrate the relationship between expected fines and reputational sanctions in this model. It is interesting to note in addition, however, that increasing fines from zero increases the reputational sanction from non-compliance at first, as just shown, but eventually reduces the reputational sanction. This is because when fines become very high, a pooling equilibrium in which all firms comply replaces the separating equilibrium. In the

compliance pooling equilibrium, the reputational sanction is no larger than the difference between the profits of a firm of unknown type and the profits of a bad firm:  $\Pi^P - \Pi^B$ . This is smaller than the sanction in the separating equilibrium,  $\Pi^G - \Pi^B$ . Increasing the legal sanction may increase or decrease the reputational sanction.

The intuition is that the reputational sanction is determined by comparing a compliant firm with a non-compliant firm. Increasing the legal sanction tends to suggest that non-compliant firms are more likely to be bad, but, if the sanction is high enough, it may also imply that compliant firms are more likely to be bad. The sanction affects the observers' updated beliefs about both compliant and non-compliant firms, and the reputational gap between them may grow or shrink as legal sanctions increase. Increasing fines affects the reputational sanction, but not in a monotonic way.

The model assumes a dichotomy of firm types: the firm is either good or bad, and there is a discrete difference between the two. The dichotomy assumption is useful to make as an expositional matter, but is not necessary to the argument. Rather, the key insight is that the reputational sanction for wrongdoing may vary depending on the level of the penalty, which would be true even if there were a continuum of types (as long as the types were not distributed across the population uniformly).

For example, suppose types are distributed normally across the population. Consider the "marginal" type for a given penalty, which is the type whose compliance costs are such that it is indifferent between complying and not complying at that penalty. Compare three cases. In the High Extreme Case, because of a very high expected legal penalty, the marginal type is an outlier that is far to the right in the distribution with very high costs of compliance. In the Average Case, the fine is lower and the marginal type is right in the middle of the distribution with

average costs of compliance. Finally, in the Low Extreme Case, the expected fine is very low and the marginal type is an outlier far to the left of the distribution.

Consider how the reputational sanction changes as the fine changes. When fines are zero or very low, the law provides little deterrence and Low Extreme Case obtains. In this case, the reputational sanction is likely to be significant, as only extremely good firms comply, while those that do not will be perceived to be close to average. As the fine increases, the Average Case will emerge at some point in which a compliant firm will be perceived to be the average of the left side of the distribution of costs, while a non-compliant firm will be the average of the right side of the distribution. The reputational sanction is larger in the Low Extreme Case than the Average Case given that the distance to the mean from an outlier in a normal distribution will be greater than the distance between the average of the left side of the distribution and the average of the right side of the distribution. Increasing the fine initially reduces the likely reputational sanction.

Further increases in the fine, however, increase the reputational sanction. As the fine increases, the High Extreme Case will emerge in which a firm that does not comply will be perceived to be a very bad firm with extremely high costs, since only very high cost firms do not comply, while firms that do comply will be seen as close to average since almost all firms comply. This creates a large reputational sanction, one that is greater than that that results in the Average Case. Thus, the reputational penalty depends on the level of the legal sanction, and the argument set out above holds even with a continuum of types. The specific relationship between fines and reputational sanctions depends on the distribution of firm types. In the two-type case, increasing the sanction increases and then reduces the sanction, while in the case of a normal distribution of types, increasing the sanction reduces and then increases the sanction.

### III. ASYMMETRIC INFORMATION AND OPTIMAL DETERRENCE

I have shown that the size of a legal penalty will affect the size of a reputational penalty. It is thus too simplistic to recommend adjusting legal penalties to account for reputational penalties without considering how reputational penalties depend on legal penalties. In this section I consider an optimal approach to legal penalties in light of their impact on reputation. There are two points to be made. First, setting aside the social value of signalling, the social planner must account for the dependence of the reputational sanction on the legal sanction when setting penalties for deterrence reasons. Second, to the extent that the signal sent by wrongdoing, or the absence of wrongdoing, creates social value, the planner must also account for these benefits when setting the legal penalty.

On the first point, even if deterrence is the only concern, simple substitution of reputational sanction for a legal sanction is generally inappropriate, since a change in the legal sanction will affect the reputational sanction. It is too simple to suggest, as Cooter and Porat (2001) do, that a given legal fine can be lowered to reflect the reputational sanction at that given level of fine; lowering the fine changes the reputational sanction. The social planner must account for endogeneity when setting a penalty to achieve optimal deterrence. That is, the legal fine plus the reputational sanction must deter optimally, and the legal fine must give rise to the signalling (or screening<sup>10</sup>) equilibrium that results in the appropriate reputational sanction.

Second, a complex problem becomes even more complex if there is a social value to resolving asymmetric information through the signalling associated with a legal penalty. In

---

<sup>10</sup> Signals are sent by the informed actor (the would-be wrongdoer in this context) while screening is done by uninformed actors. I consider in this section how lawmakers can set sanctions optimally so as to allow potential wrongdoers to signal their types. That is, uninformed actors, lawmakers, are making decisions that allow informed actors to signal their type. This could be conceived as a kind of screening, but for consistency with the previous analysis, which has in effect taken the law as given (and thus removed lawmakers from consideration), I will continue to refer to this as signalling.

general, optimal deterrence and optimal signalling cannot be simultaneously achieved with a fine. The law has one instrument, the expected legal penalty, that affects both deterrence and signalling.<sup>11</sup> It cannot be calibrated to achieve both optimal deterrence and signalling perfectly. While there remains a welfare-maximizing approach to setting the penalty, it does not give rise to the ideal outcome.

To develop this argument, first consider the social costs and benefits of signalling in the context developed above. Suppose that conduct in the first period is subject to legal penalty, but in the second period conduct is not governed by law and depends only on the firm's type. Following the firm's decision to commit a legal wrong or not in the first period, observers, depending on the nature of the equilibrium, will either gain no new information about firm type (either of the pooling equilibria) or will learn the firm's type perfectly (separating equilibrium). The firm's reputation following the first period will affect decisions and social welfare in the second period. The degree to which social welfare is affected by the revelation of firm type, however, may vary widely.<sup>12</sup>

For example, suppose that known good firms earn a higher profit in the second period than other firms not because they are able to engage in more value-added transactions, but because they are in a stronger bargaining position with the buyers that they would deal with in any event. For instance, having a good reputation may make a sale to a third party relatively more profitable for the firm, and this improved outside option increases the surplus that it can extract in its bargaining with the specific buyer. In such a case, signalling in the first period

---

<sup>11</sup> In future work I intend to explore how the law may establish nuanced penalties that attempt to account for reputation. For example, a no-contest plea results in a financial penalty, but a dampened reputational penalty relative to an acknowledgement of guilt given that there is a noisy signal about whether the actor actually committed the wrong. A declaratory judgment, on the other hand, establishes better the existence of certain conduct, but does not involve a financial penalty.

<sup>12</sup> Baniak and Grajzl (2013) also observe that the social costs of reputational sanctions vary widely, but because they focus on concerted punishment that is harmful to the punisher, do not account for the social benefits of reputational losses to the wrongdoer: observers avoid future harms because of better information about the wrongdoer's type.



creates no social benefits in the second; it simply shifts surplus between the parties. The drop in profits from a bad reputation are privately significant to the firm, but are socially insignificant.

On the other hand, a better reputation for good firms, and a corresponding lesser reputation for bad firms, might allow some value-added transactions to take place that would not otherwise. Suppose, for example, that good firms do not find it costly to provide a high quality product to a particular buyer in the second period and hence can be relied upon to produce such a product. Bad firms, on the other hand, would produce a low quality product. Moreover, suppose that the buyer will not buy at all unless it knows that the firm is good and hence that the product is high quality. In such a case, a separating equilibrium in the first period creates social surplus in the second that would not exist otherwise.

In general, social surplus is higher in the first period where the wrongdoer internalizes fully the net social cost of the first-period wrong, and is higher in the second period where activity in the first resolves asymmetric information problems about the wrongdoer's type. Existing literature recognizes the role of law and reputational sanctions in realizing the first period goal, but not the second. The problem is that the same legal instrument, expected penalties in the first period, affects deterrence in the first period and asymmetric information in the second, and thus affects surplus in different ways in different periods such that the optimum in each period is not generally available.

If both deterrence and optimal signalling are not simultaneously possible, the social planner should compare the relative benefits of deterrence and signalling and set the expected penalty to induce the preferable outcome. The social planner would have to figure out what types should optimally comply and which should not and then set a penalty accordingly. This

may lead the planner to optimally set the penalty at a point where the fine plus the reputational sanction could be higher or lower than the social cost of the wrong.

The analysis in this section generates some scepticism about the capacity of legal authorities optimally to account for reputational effects. Even under Cooter and Porat's (2001) analysis in which the reputational sanction is exogenous, adjusting the legal penalty to account for reputational effects is complex. This analysis shows that it is in reality even more complicated. The social planner would require information on, amongst other things, the likelihood of different signaling equilibria, and the social benefits of more granular information about the wrongdoer's type. There are a host of nuances in the optimal legal approach that arise when the endogeneity of optimal legal and reputational penalties is recognized.

#### IV. THE PROBABILITY OF DETECTION VERSUS FINES

To this point I have considered the expected penalty from the action and have not discussed the constituent elements of the expected penalty, the probability of detection and the size of the fine. It has been unnecessary to do so because of the simplifying assumption that observers know whether a wrong has been committed regardless of whether or not the law detects the wrong and imposes a punishment. It may be, however, that observers do not know whether a wrong has been committed unless there is an imposition of legal sanctions. I consider the implications of this possibility in this section.

Becker (1968) observed that the expected penalty for a crime is not only a function of the penalty that is imposed upon conviction, but also on the probability of detection. Given that detection is socially costly in that resources must be devoted to identifying wrongful behaviour, while fines are simply a transfer, there is an argument in favour of keeping the probability of detection low and the fine high. There are a number of responses to this argument, such as the

importance of marginal deterrence (Stigler, 1970), and risk aversion (Polinsky and Shavell, 1979). The analysis in this article provides another reason to keep the probability of detection high.<sup>13</sup>

In the framework set out above, if observers are only aware of a wrong in the event of legal detection, then the information that observers have about firm type following the first period will generally not be as precise even where there is a separating equilibrium. If the wrong is detected and the firm punished, then in a separating equilibrium, observers are certain of the firm's type. But if the firm is not punished, there are two possibilities: the firm is good and did not commit the wrong; or the firm is bad but was not detected committing the wrong. The absence of a penalty increases the probability that the firm is good, but the probability does not, as it did in the above discussion, go to one. Even if the reputational and legal penalties are such that good firms comply and bad firms do not, which is possible even with a low probability of detection if the fine is high enough, observers will not be certain of type in the second period unless a wrong is detected and punished in the first.<sup>14</sup>

To elaborate, if there is a separating equilibrium, observers would follow Bayes' Rule to identify the probability that a firm that was not punished in period one is a good firm. Let the prior probability of a good firm be  $\gamma$ . An unpunished firm may be good, or it may be bad but fortuitously for it unpunished. The probability of a good firm conditional on no punishment in period one is the fraction of good firms over the sum of the fraction of good firms and the

expected fraction of bad firms that go unpunished:  $\alpha = \frac{\gamma}{\gamma + (1-\gamma)(1-p)}$ .

---

<sup>13</sup> The analysis resonates with that of Baker and Malani (2011) who suggest that a trial process that discourages settlement may be socially advantageous by better publicizing possible wrongs. Posner (2000) and Shapira (2013) also emphasize the role of legal proceedings in providing information that allows informal sanctions to arise.

<sup>14</sup> Similarly, if authorities sometimes erroneously punish a non-wrongdoer, there will be noise about type even in the presence of a fine. For simplicity, I focus in the text on the case of non-detection.

Where the probability of detection is very low, there will be little updating of the observers' prior beliefs about firm type when there is no punishment; as the probability of detection,  $p$ , goes to zero, the posterior belief about firm type,  $\alpha$ , will approach the anterior belief,  $\gamma$ . Even bad firms are rarely detected, so there is little to distinguish between types ex post even if good firms comply and bad firms do not.

Even though there could be a separating equilibrium in which only good firms comply and bad firms do not, the separation has very little impact on reputation if the probability of detection is low. If the probability that a firm is good, conditional on not being punished, is relatively low since many bad firms go unpunished, then the separation of types is not very informative to observers. Assuming that the more precision observers have about firm type, the greater the welfare gains from resolving asymmetric information, a low probability of detection has negative implications for welfare in the second period.

This analysis suggests a reason to maintain a significant probability of detection of wrongs even if doing so creates social costs. By maintaining a significant detection rate, noise that affects signals from the (non-)imposition of first period legal penalties is lessened, which improves information about firm type in later periods. Better information will, in general, lead to welfare gains.

## V. CONCLUSION

The analysis has shown that reputational sanctions are not independent of legal penalties, but rather in part depend on them. Changing legal penalties to account for the reputational effects of misconduct changes the reputational effects. The analysis suggests that optimizing

legal penalties in light of reputational effects is even more complex than the literature recognizes.

Signalling further suggests that there is good reason from a social perspective to maintain significant probabilities of detection of wrongs. Without such a probability, if observers cannot themselves easily observe whether a wrong has been committed, good firms will resemble bad firms and the signal sent from non-punishment will be weak. A stronger probability of detection helps distinguish good firms from bad.

I conclude by relating the analysis to some aspects of existing empirical evidence on reputational sanctions. The analysis suggests two shortcomings of the empirical literature. First, existing empirical studies of public companies understate the reputational impact of misconduct. Studies generally examine the drop in market capitalization following a finding of misconduct. If the conduct is observed even without legal proceedings, or if the probability of detection is significant, the correct comparison is not between market capitalization before and after, but rather between what market capitalization would have been with the better reputation associated with good conduct and what the market capitalization was after bad conduct.

Second, some studies that have examined the interaction between legal sanctions and reputational sanctions provide results that are in some ways incomplete. Armour et al. (2011), for example, examine the relationship between the level of a legal penalty when levied and the associated reputational sanction. They find little connection between the two. This result seems inconsistent with an expressive function of law: if the seriousness of the sanction were viewed as an indicator of how the wrongdoer should be informally punished (see, e.g., Teichman 2005), then there should be an observed correlation between formal and informal penalties, but there is not.

In contrast, the result that the reputational sanction is not generally associated with the size of the legal penalty is consistent with the understanding of the reputational sanction in this article. As a preliminary matter, the connection between legal sanctions and reputational sanctions depends on *expected* penalties, not actual penalties. Comparing informal sanctions with penalties that are imposed in fact does not capture what expected penalties are, and thus does not test the connection between reputation and expected penalties set out in this paper. Moreover, it is plausible that higher penalties are imposed precisely where the probability of detection is especially low, following the traditional logic of Becker (1968). If so, then high ex post penalties may systematically overstate expected penalties, and low ex post penalties would understate expected penalties, which would muddy the relationship between penalties actually imposed and expected penalties.

In any event, as explained above, the theory in this paper shows that there is a connection between expected legal penalties and informal reputational sanctions, but it does not necessarily imply that the higher are legal penalties, the higher are reputational sanctions. The reputational sanction properly conceived is not based on the difference in reputation before and after the wrong, but rather derives from reputation after the wrong compared to reputation if the wrong were not committed. Increasing the legal sanction affects the population of firms that commit the wrong, but of course also changes the population of firms that do not commit the wrong. Increasing the legal sanction would suggest the firms that still commit the wrong are worse, but increasing the legal sanction will also tend to imply that firms that do not commit the wrong are also worse. Either effect may dominate.

## REFERENCES

- Alexander, Cindy. 1999. On the Nature of The Reputational Penalty for Corporate Crime: Evidence. *Journal of Law and Economics* 42:489-526.
- Armour, John, Colin Mayer and Andrea Polo. 2010. Regulatory Sanctions and Reputational Damage in Financial Markets. CEPR Discussion Paper No. DP8058. online: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1711071](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1711071)
- Baniak, Andrzej and Peter Grajzl. 2013. Equilibrium and Welfare in a Model of Torts with Industry Reputation Effects. SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2038072](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2038072).
- Baker, Scott and Anup Malani, "Does Accuracy Improve the Information Value of Trials?" (2011) [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1945421](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1945421).
- Becker, Gary. 1968. Crime and Punishment: An Economic Approach". *The Journal of Political Economy* 76:169–217.
- Bernstein, Lisa. 1992. *Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry*. *Journal of Legal Studies* 21:115-157.
- Bierschbach, Richard A. and Alex Stein. 2004-2005. Overenforcement. *Georgetown Law Journal* 93: 1743-1781.
- Cooter, Robert and Ariel Porat. 2001. Should Courts Deduct Nonlegal Sanctions From Damages? *Journal of Legal Studies* 30: 401-422.
- Cooter, Robert. 2000a. Do Good Laws Make Good Citizens? An Economic Analysis. *Virginia Law Review* 86:1577-1601.
- Cooter, Robert. 2000b. Three Effects of Social Norms on Law: Expression, Deterrence, and Internalization. *Oregon Law Review* 79: 1-22.
- Ganunza, Juan-Jose, Fernando Gomez-Pomar and Marta Robles. 2011. Product Liability versus Reputation. Available at SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1986237](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1986237).
- Ellickson, Robert. C. 1991. *Order Without Law: How Neighbors Settle Disputes*. Cambridge, Mass: Harvard University Press.
- Iacobucci, Edward. 2006. Predatory Pricing, the Theory of the Firm and the Recoupment Test: An Examination of Recent Developments in Canadian Predatory Pricing Law. *The Antitrust Bulletin* 51: 281-323.
- Kahan, Dan. 1997. Social Influence, Social Meaning, and Deterrence. *Virginia Law Review* 83: 349-395.

- Kahan, Dan M. and Eric A. Posner. 1999. Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines. *Journal of Law and Economics* 42: 365-392.
- Karpoff, Jonathan and John Lott. 1993. The Reputational Penalty Firms Bear From Committing Criminal Fraud. *Journal of Law and Economics* 36: 757-802.
- Karpoff, Jonathan M., John R. Lott, Jr., and Eric W. Wehrly, 2005. The Reputational Penalties for Environmental Violations: Empirical Evidence. *Journal of Law & Economics* 48:653-675.
- Karpoff, Jonathan M., Scott D. Lee and Gerald S. Martin. 2008. The Cost to Firms of Cooking the Books. *Journal of Financial and Quantitative Analysis* 43: 581-611.
- Klein, Benjamin, and Keith Leffler. 1981. The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy* 89:615-641.
- Lessig, Lawrence. 1995. The Regulation of Social Meaning. *The University of Chicago Law Review* 62: 943-1045.
- Macaulay, Stewart. 1963. Non-Contractual Relations in Business: A Preliminary Study. *American Sociological Review* 28: 55-67.
- McAdams, Richard. 2000. A Focal Point Theory of Expressive Law. *Virginia Law Review* 86: 1649-1729.
- Mitchell, Mark and Michael Maloney. 1989. Crisis in the Cockpit? The Role of Market Forces in Promoting Air Travel Safety. *Journal of Law and Economics* 32: 329-355.
- Murphy, Deborah L., Ronald E. Shrieves and Samuel L. Tibbs. 2009. Understanding the Penalties Associated with Corporate Misconduct: An Empirical Examination of Earnings and Risk. *Journal of Financial and Quantitative Analysis* 44: 55-83.
- Peltzman, Sam. 1981. The Effects of FTC Advertising Regulation. *Journal of Law and Economics* 24: 403-448.
- Polinsky, A. Mitchell and Steven Shavell. 2010. The Uneasy Case for Product Liability. *Harvard Law Review* 123:1437-1492.
- Polinsky, A. Mitchell and Steven Shavell. 1979. The Optimal Tradeoff between the Probability and Magnitude of Fines. *The American Economic Review* 69: 880-891.
- Posner, Eric. 2000. *Law and Social Norms*. Cambridge, Mass: First Harvard University Press.
- Scott, Robert. 2000. The Limits of Behavioral Theories of Law and Social Norms. *Virginia Law*



*Review* 86:1603-1647.

Shapira, Ray. (2013). Reputational Sanctions in the Shadow of the Law.

Stigler, George. 1970. The Optimum Enforcement of Laws. *The Journal of Political Economy* 78: 526-538.

Teichman, Doron. 2005. Sex, Shame, and the Law: An Economic Perspective on Megan's Laws. *Harvard Journal on Legislation* 42:355-415.