

2012

## **Don't Say You're Sorry Unless You Mean It: Pricing Apologies to Achieve Credibility**

Murat C. Mungan

*Florida State University College of Law*

Follow this and additional works at: <https://ir.law.fsu.edu/articles>



Part of the [Criminal Law Commons](#), and the [Law and Economics Commons](#)

---

### **Recommended Citation**

Murat C. Mungan, *Don't Say You're Sorry Unless You Mean It: Pricing Apologies to Achieve Credibility*, 32 *INT'L REV. L. & ECON.* 178 (2012),

Available at: <https://ir.law.fsu.edu/articles/116>

This Article is brought to you for free and open access by Scholarship Repository. It has been accepted for inclusion in Scholarly Publications by an authorized administrator of Scholarship Repository. For more information, please contact [efarrell@law.fsu.edu](mailto:efarrell@law.fsu.edu).

# Don't Say You're Sorry Unless You Mean It: Pricing Apologies to Achieve Credibility

Murat C. Mungan\*

Florida State University College of Law

November 23, 2011

## Abstract

Remorse and apologies by offenders have not been rigorously analyzed in the law and economics literature. This is perhaps because apologies are regarded as '*cheap talk*' and are deemed to be non-informative of an individual's conscious state. In this paper, I develop a formal framework in which one can analyze remorse and apologies. I argue that legal procedures can be designed to *price* apologies, such that only truly remorseful individuals apologize. Hence, apologies would not be mere '*cheap talk*' and could send correct signals regarding an offender's true conscious state, making them credible. This will lead victims, upon receiving apologies, to forgive offenders more frequently. Moreover, pricing apologies does not negatively impact the possibility of achieving optimal deterrence. An (arguably negative) effect of pricing apologies is its elimination of insincere apologies. If it is assumed that apologies, even if insincere, carry rehabilitative and/or palliative benefits, then the

---

\*Visiting Assistant Professor of Law, Florida State University College of Law, 425 W. Jefferson Street, Tallahassee, FL 32306. e-mail: mmungan@law.fsu.edu. I would like to thank two anonymous referees whose comments lead to significant improvements in the content and organization of this article. I also thank D. Bruce Johnsen for useful suggestions. The usual disclaimer applies.

optimality of pricing apologies depends on a trade-off between achieving credibility and increasing such rehabilitative and palliative benefits.

**Keywords:** Apology, Crime and Deterrence, Remorse, Optimal Sanctions.

## 1. Introduction

Unless apologies are associated with a form of negative consequence, they must be regarded as *cheap talk*.<sup>1</sup> In criminal law apologies are associated with expected positive rather than negative consequences. In most legal systems, offenders who convincingly assert that they are remorseful are given a reduction in penalty.<sup>2</sup> In the United States, in many jurisdictions judges have discretion in determining the offender's sentence, who may be affected by whether the offender displayed remorse and how convincingly he has done so.<sup>3</sup>

In the initial analysis, from a utilitarian perspective, the treatment of remorse in this fashion makes little sense. In most utilitarian approaches, penalties are imposed to achieve efficient deterrence. Hence the only thing that is relevant is the expected magnitude of penalties. As long as offenders react to incentives, they will consider the likely punishment that they will get, and decide whether they should commit the offense. Imposing lower sanctions on those who

---

<sup>1</sup>See Frank (2005 p. 25) stating that apologies are viewed as cheap talk in self-interest models. He states: "The traditional law and economics model asks us to interpret laws and regulations as means that enable narrowly self-serving actors to better achieve their ends. Yet numerous details of legal custom and practice appear inconsistent with this interpretation." Frank (2005 p. 25). He then lists the inclusion of remorse as a factor in sentencing as one such detail. My model can be interpreted as suggesting that the inconsistency identified by Frank is not the result of assuming 'narrowly self-serving actors'.

<sup>2</sup>See Frank (2005 p. 25).

<sup>3</sup>See, for instance, Federal Rule of Criminal Procedure 32(i)(4), which outlines the procedure for what is commonly called *allocution*. In particular, Rule 32(i)(4) states:

"Before imposing sentence, the court must:

...

(ii) address the defendant personally in order to permit the defendant to speak or present any information to mitigate the sentence"

See also O'Hear (1997) for a review of how remorse is considered in federal courts.

convincingly display remorse will therefore have an effect of lowering the expected punishment. But that level of expected punishment can easily be achieved by choosing a single sanction and imposing it on every offender, regardless of how remorseful they are. So why don't we simply impose uniform sanctions?

It has been argued, that the reason may be associated with giving offenders an incentive to apologize, which would in turn relieve the offender and the victim from certain psychological burdens.<sup>4</sup> This goal (relieving those involved from psychological burdens) is one that should be accounted for in a utilitarian social calculus when making normative statements concerning legal regimes. However, legal regimes may not be providing the correct means to achieve this legitimate goal. An offender has nothing to lose by apologizing, because doing so can decrease but not increase his penalty.<sup>5</sup> Accordingly, many non-remorseful offenders will take the opportunity and apologize. Hence, the legal regime will not create credible apologies, and victims will be suspicious about the true motives of offenders.

Proponents of current legal procedures may argue that decision makers can partially distinguish between truly remorseful and non-remorseful individuals, and that on average more

---

<sup>4</sup>See Bibas and Bierschbach (2004 p. 91), arguing that "Remorse and apology neither displace nor justify punishment, but, as functions of punishment, they can better complement and serve its goals." and that "Plea and sentencing procedures should include explicit roles for remorse and apology."

See also Ward (2006 p. 139), summarizing views expressed in several articles (for instance Pipes and Alessi (1999), Tieger (2003), Simons (2003), and Sarat (1999)), as to why remorse is or should be a relevant concept in criminal procedure.

<sup>5</sup>One may argue that apologizing may increase the *probability* of conviction, and that therefore offenders incur expected losses by apologizing. This argument overlooks the fact that a confession does not necessarily imply an apology. It is the confession that increases the probability of conviction, not the apology. Confessions and apologies need not be bundled together in legal systems. Self-reporting, for instance, is a procedural mechanism which enables confessions, but does not require an apology. Accordingly, it may make sense to decrease penalties for self-reporting, but this does not imply that similar reductions must be made for apologizing. See section 5.B. for a discussion of how procedural mechanisms such as self-reporting affect my analysis.

remorseful individuals will benefit from discounted sanctions. Although this argument is somewhat convincing, it overlooks an important problem: "sort[ing] out the truly remorseful defendant from the unrepentant but savvy defendant." (Ward (2006 p. 164)) The current legal system rewards those who can convincingly pretend as if they are remorseful, even if they do not feel the least sorry about their actions. This is problematic even if one leaves aside problems associated with creating distrust and anger on the victim's side. It is quite plausible to think that asymmetric ex-ante incentives would be provided to those who know they can fake remorse, and those who cannot. This asymmetry will necessarily cause problems of under-deterrence as well as over-deterrence.<sup>6</sup> Hence, even from a purely utilitarian stand-point, current procedures are problematic.

These observations lead one to think about whether there is an alternative.<sup>7</sup> Given the objective of relieving offenders and victims from psychological burdens, can one design legal procedures, where offenders can credibly apologize to their victims and seek forgiveness? I argue that this can be achieved. By *pricing* apologies, one can separate remorseful individuals from non-remorseful offenders. Such pricing can be obtained by slightly increasing the punishment for individuals who elect to apologize. In the remaining parts of the paper, this procedure will be called *pricing* and the increase in the penalty will be called the *price* of apologizing.

Separating remorseful and non-remorseful offenders by pricing apologies is possible, because remorseful individuals will presumably have something to gain by apologizing, namely a relieved conscience, whereas non-remorseful offenders will not.<sup>8</sup> Therefore, given increased expected

---

<sup>6</sup>Problems associated with under-deterrence and over-deterrence are common when there is an unobservable heterogeneity among potential offenders. See for instance Shavell (1990) and Mungan (2011) where the asymmetry is caused by individuals' unobservable talent levels.

<sup>7</sup>See Ward (2006) arguing that remorse should not be considered at all in criminal procedure.

<sup>8</sup>See Mungan (2010) formalizing the idea that an offender may feel remorse after committing crime and that he may partially relieve his conscience if convicted. See also Adler (1992), Bernick (1982), Tasioulas (2007),

costs associated with apologizing, non-remorseful offenders would never apologize.<sup>9</sup> On the other hand, if the increase in expected sanctions are sufficiently low, remorseful offenders will find it in their interest to apologize and relieve their guilty conscience. Hence, pricing apologies should suffice to distinguish remorseful offenders from those who are not.

Separation of remorseful and non-remorseful individuals, through pricing apologies, is beneficial due to a variety of reasons. Separation will lead to a credible apology system, where offenders can signal to their victims that they are truly sorry about their actions. This will not only relieve part of the offender's guilty conscience but also allow the victim or his family to be certain of the offender's true motive.<sup>10</sup> This will lead to a higher frequency of forgiveness, which would mitigate the victim's psychological burdens arising from hate towards the offender.<sup>11</sup> Moreover, as I will demonstrate in the proceeding sections, pricing apologies does not negatively impact the possibility of achieving optimal-deterrence. Accordingly, attaching a price to apologizing leads to benefits by fixing a problem in the current legal procedures, namely the non-credibility of apologies by offenders. Thus, from a utilitarian perspective, pricing apologies is appealing.

My argument, however, is still incomplete. It overlooks the fact that sanctions can be costly to impose. This is especially true for severe crimes, where optimal deterrence requires incarceration.<sup>12</sup> For these crimes, a sentence reduction will lead to savings from the cost of

---

Cochran (1998) and Bibas and Bierschbach (2004).

<sup>9</sup>If there are external benefits from apologizing, even non-remorseful individuals may wish to apologize. Incorporating external benefits would imply that the price of apologies must be increased. Section 5.D. extends the analysis to cases where external benefits are present.

<sup>10</sup>Bibas and Bierschbach (2004) mentions similar gains associated with manifestations of remorse.

<sup>11</sup>See Bibas and Bierschbach (2004) for a more extensive list of benefits associated with manifestations of remorse.

<sup>12</sup>See Polinsky and Shavell (1984), Shavell (1987) and Kaplow (1990) analyzing optimal non-monetary sanctions.

incarceration. Hence, one could argue that decreasing the penalty for remorseful offenders could lead to substantial savings. This argument, however, ignores the problems that I have sought to solve in the first place. Reducing penalties for apologizing makes it impossible to perfectly separate remorseful offenders from those who are not remorseful. Hence, it is not possible to decrease the penalty for remorseful offenders only. Therefore, penalty reductions for (convincingly) apologizing will come with costs associated with the problems I have identified earlier. Whether reductions in incarceration costs will outweigh costs associated with a non-credible apology system is not clear. Furthermore, if incarceration costs were substantial, one could uniformly decrease the penalties for remorseful as well as non-remorseful individuals in a priced apology system. This solution would result in similar savings, without producing costs associated with a non-credible apology system. I provide a more detailed and formal discussion of this issue in section 5, after I provide the necessary tools for interpretation in sections 2-4.

It should also be noted that the existence of other procedural incentive mechanisms, such as self-reporting and plea bargaining, have no effect on the validity of what I am claiming.<sup>13</sup> Here, I do not enter the debate about if and how these procedural tools should be used. I am simply stating that, given any punishment scheme for self-reporters, plea-bargainers and those who go to trial, a slightly increased penalty scheme should apply if the person desires to apologize. Penalties for apologizing are aimed completely at distinguishing remorseful individuals from non-remorseful offenders, thereby creating credibility. Increased penalties, for those who wish to apologize, can be designed so as to not affect the ex-ante incentives of individuals to commit crime. As such, benefits associated with priced apologies continue to exist, without affecting the objectives that procedural tools, such as self-reporting or plea bargaining, serve.

---

<sup>13</sup>There is a very broad and expanding literature on self-reporting and plea bargaining. A few articles analyzing self-reporting are Kaplow and Shavell (1994), Innes (1999), (2000) and (2001), and Mungan (2011).

One may argue, from a non-utilitarian perspective, that pricing apologies is in tension with our intuition that remorseful offenders should be punished less severely. This intuition is presumably guided by our sympathy towards a person who has realized his wrongs. If so, we must re-evaluate our intuitions, because, under a priced apology system, a remorseful offender is better off by apologizing. In fact, no offender is forced to apologize and be subject to increased penalties. The system merely provides the opportunity to apologize, and those who are truly remorseful voluntarily do so. Hence, being sympathetic towards remorseful offenders should push us towards being more supportive of pricing apologies.

Another objection to pricing apologies may come from a retributive approach. A retributivist may feel that a person who feels remorse upon committing crime is now a *changed man*. Accordingly, we are not punishing the person who has committed crime, but a different person who does not deserve to be punished. The implication of this reasoning would be that remorseful individuals should not be punished at all, as opposed to being punished with a reduced sentence.<sup>14</sup> However, achieving this result is impossible, unless we are ready to let all criminals walk away without punishment. This follows from the impossibility of simultaneously (i) reducing penalties for remorseful offenders and (ii) separating remorseful from non-remorseful offenders. Hence, the only way to not punish remorseful offenders would be to not punish any offender. Therefore, addressing these objections in a reasonable manner is impossible, not only under a priced apology system, but in any system.<sup>15</sup>

---

<sup>14</sup>One can formulate a consequentialist and milder variant of this argument. The unit cost of punishing a truly remorseful offender may be smaller than punishing a non-remorseful offender, because the former may be less likely to recidivate. This suggests that it is still ideal to punish truly remorseful offenders, but *less severely* than non-remorseful offenders. But this argument has a similar caveat: It is impossible to simultaneously (i) reduce penalties for remorseful offenders and (ii) separate them from non-remorseful offenders.

<sup>15</sup>For another view as to why remorse should not be relevant from a retributivist perspective, see Ward (2006).

A stronger objection to pricing apologies can be formed by emphasizing potential social gains from insincere apologies. Apologies of all types -credible and non-credible- may produce some rehabilitative and palliative benefits. Apologies may assist in reforming the attitudes of an offender towards crime and society, they may help in rehabilitating offenders. They may also have a palliative effect for victims. Victims may derive pleasure from seeing offenders humiliated or humbled. It is true that insincere apologies may generate such benefits, but it is equally plausible to think that they may generate some additional social costs. These include creating additional distrust and anger on the victim's side, and also generating asymmetric incentives for individuals who can fake remorse and for those who cannot (which creates costs of under and over deterrence). As such, the size and sign of the effects of insincere apologies are debatable. Nevertheless, in sub-section 5.E., I provide an analysis assuming that insincere apologies lead to net social benefits. In this case, it is not possible to immediately conclude that pricing regimes are optimal, but one can determine conditions under which they are. Unsurprisingly, pricing regimes are likely to be optimal when the value of credibility is high or the net social benefits of insincere apologies are low. But a second -and perhaps counter-intuitive- result is revealed. Pricing regimes are more likely to be optimal if the proportion of remorseful offenders is high. The reason is that a pricing regime is one that enables the truly remorseful offender to psychologically relieve his victim by making credible apologies possible. As such, the value of such a regime is naturally increasing in the proportion of remorseful offenders.

To the best of my knowledge there has not been many attempts in modeling remorse in the law and economics literature.<sup>16</sup> This may be because, absent increased penalties for apol-

---

<sup>16</sup>An exception is Mungan (2010). However, this article does not analyze apologies. As such, it implicitly assumes that apologies are *free*. Accordingly, one could interpret this article as analyzing the case where

ogizing, economists are tempted to view manifestations of remorse as *cheap talk*. Furthermore, at first glance, one may conjecture that no rational individual would apologize if it was costly to do so. Hence, law and economics scholars may have concluded that remorse is not worth modeling, because it should not be relevant in determining the offender's sentence. Since I reach the opposite conclusion, in the proceeding parts of this paper, I develop an economic optimal deterrence model to verify my claims.

It is also worth noting that the instant paper can be placed among several recent articles in regulatory theory that either analyze how costly signals can be (or are) used to extract information about regulatees' privately known and heterogenous characteristics. For instance, Raskolnikov (2009) proposes a method to separate tax payers, who game the system<sup>17</sup> from those who "comply out of habit, a sense of duty or reciprocity, a desire to avoid feelings of guilt or shame, and for many other reasons"<sup>18</sup>. Similarly, Stephenson (2006) discusses how and when high quality, and therefore costly, explanations for decisions made by agencies and legislators can serve as signals that inform judges of how beneficial these decisions will be for society. Toffel and Short (2011) can also be counted among examples of articles that demonstrate how costly signals can be used in designing welfare enhancing regulatory schemes.

In the next section, I challenge the proposition that no rational individual would apologize if by doing so they would face higher expected sanctions. I rely on previous observations by scholars of law concerning the value of seeking forgiveness to offenders, who regret their previous actions and seek atonement. By incorporating these observations, I model a variant of the standard 'rational criminal' as depicted by Becker (1968), and describe the framework remorseful offenders are presumed to apologize when convicted and relieve their guilty conscience.

<sup>17</sup>In this article one games the system if he is "paying only when the cost of noncompliance outweighs its benefits.", see Raskolnikov (2009) at p. 689.

<sup>18</sup>*Id.*

I use to represent and formalize 'remorse' in the context of optimal deterrence models. This approach provides a formal model, in which remorseful individuals are willing to pay a price for apologizing, and provides a rationale as to why methods enabling credible apologies should be made available.

In section 3, I derive a utilitarian social welfare function to evaluate the desirability of various sanction schemes by incorporating the framework designed to formalize remorse in section 2. Formalizing a method to compare different policies in this manner enables the identification of costs and benefits associated with various treatments of remorse. These include standard costs and benefits associated with deterrence and benefits from (partially) relieving offenders and victims from psychological burdens through credible apologies. Furthermore, comparing these costs and benefits allows the determination of optimality conditions. These conditions are derived in section 4, where I demonstrate that priced apologies are optimal. In section 5, I discuss various extensions and implications of my model, and I conclude in section 6. An Appendix contains proofs of various propositions.

## **2. Remorse and Apologies**

Standard models of crime and deterrence ignore the fact that criminals may regret their previous behavior and feel remorse.<sup>19</sup> However, remorse is an important concept, and policy implications derived from models which ignore its existence may suggest sub-optimal legal procedures. Furthermore, standard models which do not account for remorse can fail to explain observed human behavior. To address these problems, I provide an illustrative crime and deterrence model that incorporates remorse. It should be noted, however, that this model does not purport to offer a complete analysis of optimal sanctioning schemes for apologies. This

---

<sup>19</sup>See Garoupa (1997) and Polinsky and Shavell (2000) for a review of the existing economics literature on law enforcement.

model does not, for instance, allow us to determine the optimal method of pricing apologies.<sup>20</sup>

These issues are instead discussed informally in sub-section 5.F.

Consider an individual, who turns him-self in for murder many years after committing it, despite being confident that there is no evidence against him.<sup>21</sup> Call this person *Otto*.<sup>22</sup> In standard crime and deterrence models, *Otto* cannot exist. This is because *Otto* is not allowed to have benefits associated with self-reporting. To be specific, in standard crime and deterrence models, a potential offender's preferences are represented by the following utility function:

$$U = \begin{cases} 0 & \text{if he does not commit crime} \\ b & \text{if he commits crime and is not caught} \\ b - s & \text{if he commits crime and is caught} \end{cases} \quad (1)$$

Where  $b$  is the material benefit from crime, and  $s$  is the punishment associated with that crime. Hence, by self-reporting, an individual is incurring a cost of  $s$ , but is not gaining anything.<sup>23</sup>

However, in reality, an individual who is remorseful, may desire to turn himself in to relieve his guilty conscience. This type of behavior has been observed, documented and discussed by various scholars of law.<sup>24</sup> A model which allows for individuals like *Otto*, must take at least two things into consideration: (i) the fact that individuals may experience psychological burdens after committing crime, and (ii) the fact that they seek to mitigate this burden. Next, I will propose a variant of the standard utility function which takes both points into consideration.

---

<sup>20</sup>It should also be noted that this model abstracts from issues related to the incapacitation function of imprisonment. Incorporating potential incapacitation benefits may generate additional trade-offs. See, e.g., note 14 *supra*.

<sup>21</sup>See Bernick (1982) for similar real life examples.

<sup>22</sup>Named Otto after the example in Bernick (1982).

<sup>23</sup>This is obviously true even if there is a discounted but positive penalty for self-reporting.

<sup>24</sup>See Mungan (2010) for a list of articles discussing related issues.

To incorporate the first point, namely that individuals may have psychological problems after committing crime, an important thing to note is that regret is by definition an *ex-post* concept. That is to say, one can not know how sorry one will feel, until one commits crime. Hence, in economic terms, it is best to think of remorse as a stochastic concept. Accordingly, psychic costs associated with regret can be modeled as a random variable  $G$ . The value  $G$  will take is unknown to potential offenders before they commit crime, but is realized after the commission of the crime. Hence, an offender's expected net benefits from the commission of the crime is given by:

$$E[\Pi_{c'}] = b - ps - E[G] \quad (2)$$

Where  $p$  denotes the probability of detection, and  $E[G]$  is the expected value of  $G$ . To simplify the analysis, I assume that  $G$  will take a positive value ( $g$ ) if the offender experiences remorse after committing crime, and that it will equal zero if the offender is not remorseful. Furthermore, I assume that an individual who commits crime, will feel remorse with a probability of  $\alpha$ . Incorporating these simplifying assumptions (2) becomes:

$$E[\Pi_{c'}] = b - ps - \alpha g \quad (3)$$

My modifications do not yet account for the fact that remorseful individuals may seek to mitigate psychological burdens through various activities. In general, there may be many ways to achieve such mitigation. For instant purposes, what needs to be incorporated is that remorseful offenders can mitigate psychological burdens by apologizing to their victims.<sup>25</sup> To incorporate the mitigation effect of apologies, a remorseful offender will be assumed to increase

---

<sup>25</sup>See Bibas and Bierschbach (2004 p. 116-117) for a review of studies documenting the fact that offenders feel better and happy after apologizing, and that they appreciate the opportunity to apologize. In particular, see Bibas and Bierschbach (2004 p. 117) quoting Netzig and Trenzcek (1996): "Offenders welcomed the chance to "explain their own behavior, apologize, ease their consciences and reduce feelings of guilt."

his utility by  $r > 0$ , after apologizing to his victim. Hence, a person's benefit will depend on whether or not he apologizes. Let  $A = 1$  denote an offender's choice to apologize and  $A = 0$  denote his choice to abstain from apologizing. Using this notation, a person's relief of conscience can be represented by the following:

$$R(A, G) = \begin{cases} r > 0 & \text{if } A = 1 \text{ and } G = g \\ r = 0 & \text{otherwise} \end{cases} \quad (4)$$

It should also be noted that sanctions ( $s$ ), can potentially depend on whether or not the person apologizes. The easiest way to incorporate these two modifications is to summarize the net benefits of an offender who is caught ( $\Pi_c^d$ ), and an offender who is not ( $\Pi_c^n$ ). Both expressions depend on whether or not the offender turns out to be remorseful and  $\Pi_c^d$  depends also on whether or not he apologizes<sup>26</sup>:

$$\Pi_c^d = b - s(A) - G + R(A, G) \quad (5)$$

and

$$\Pi_c^n = b - G \quad (6)$$

Eqn. (5) demonstrates the use of modifying the standard 'rational criminal'. The modifications I have made provides a framework in which remorseful individuals find it in their self-interest to apologize, even if it is costly to do so. This is summarized by the following observation.

**Observation 1:** Detected remorseful offenders choose to apologize as long as apologizing does not lead to significant penalties (i.e.  $r > s(1) - s(0)$ ).

---

<sup>26</sup>The expression for  $\Pi_c^n$  reflects my implicit assumption that individuals who are not caught cannot self-report. This is to simplify the analysis and allowing self-reporting does not affect results. See sub-section 5.B., for a discussion on the robustness of results.

To verify observation 1, note that a remorseful individual who is caught will have a benefit of  $b - s(1) - g + r$  if he apologizes, and a benefit of  $b - s(0) - g$  if he does not. Hence, he would be willing to apologize as long as  $r > s(1) - s(0)$ .

In sum, the simple modifications I made on the offender's benefit structure allow us to incorporate remorse and the value of apologizing into crime and deterrence models. These simple modifications embody my initial presumption that remorseful offenders have something to gain, namely a relieved conscience, by apologizing, whereas non-remorseful offenders do not. In the next sections, I derive optimality conditions by making use of this new structure and exploiting remorseful offenders' willingness to apologize.

### 3. Social Welfare

Given that remorseful offenders are willing to pay a price for apologizing, it should be a straightforward task to obtain separation between remorseful and non-remorseful offenders. This can be achieved by imposing slightly higher fines to individuals who apologize. However, it is not equally clear whether pricing apologies will negatively affect the possibility of achieving optimal deterrence. Furthermore, I have not yet formalized the value of credible apologies. In this section, I derive a social welfare function which accounts for three sources of costs and benefits: (i) victim's psychic costs, (ii) offender's psychic costs, and (iii) deterrence (or total harm to victims). I show that pricing apologies maximizes this function.

#### A. Victim's Psychic Costs

Victims are assumed to incur material and psychic costs totaling  $h$  and relieve part of their psychic costs upon hearing credible apologies.<sup>27</sup> To reflect these observations, I will assume

---

<sup>27</sup>See Bibas and Bierschbach (2004 p. 116) quoting Strang and Sherman (2003): "The [empirical] evidence suggests that victims see emotional reconciliation to be far more important than material or financial reparation." They also state: "Victims, offenders, and community members who have met and engaged in apologetic

that the value of an apology depends on how credible it is. To model credibility in this regard, I assume that the value of an apology to the victim will be a function of the proportion of true apologies. Let  $q$  denote the proportion of sincere apologies (accordingly,  $1 - q$  is the proportion of fake apologies). The value of an apology is denoted as  $v(q)$ . Hence, total harm is discounted by  $v(q)$  and is given by  $h - v(q)$ .

Furthermore, I assume the following intuitive properties:  $v'(q) > 0$  and  $v(\alpha) = 0$ . This property states that apologies are valuable only if they give better signals about the conscious state of an offender than no signal at all. To see this, imagine a situation where no offender apologizes. In this case, a rational victim would assume that the offender is remorseful with a probability of  $\alpha$ , reflecting the proportion of offenders who are remorseful. When  $q = \alpha$ , victims would receive the same information from an apology. In other words, an apology does not update the victim's belief concerning the conscious state of the offender. On the other hand, when  $q > \alpha$ , he knows that an apologizing offender is remorseful with a probability higher than  $\alpha$ .<sup>28</sup>

One may legitimately object to this assumption, arguing that even insincere apologies have some social value, in other words, one can claim that  $v(\alpha) > 0$  should hold. Equivalents of this assumptions are discussed in section 5.A., where I discuss victims' benefits from apologies in more detail.

## **B. Offenders' Utilities**

Offenders' utilities are described to a great extent in section 2. In particular, Equations (5) and (6) in the previous section describe an offender's benefits from committing crime. A few

---

discourse overwhelmingly feel satisfied and relieved." Bibas and Bierschbach (2004 p. 116).

<sup>28</sup>One may question whether  $v(q) < 0$  if  $q < \alpha$  should hold. Although this is an interesting question, it is not relevant for instant purposes because  $q < \alpha$  is not, and cannot be, observed in this model, because a remorseful offender always has greater incentives to apologize than a non-remorseful offender.

additional things must be formalized before proceeding further.

A potential offender, who is deterred from committing crime, has a net benefit of zero, because he is not subject to a sanction and receives no criminal benefits. Furthermore  $b$ , the benefit from crime, is assumed to vary across potential offenders. To reflect this assumption,  $f(b)$  (which is positive for  $b \in [0, \infty)$ ) denotes the density of  $b$ , and  $F(b)$  is the cumulative distribution function with  $F(\infty) = 1$ .

Another thing worth noting is that sanctions ( $s(0)$  and  $s(1)$ ) are assumed to be transferrable. Accordingly, they are costs to the offender, which appear as benefits elsewhere. Therefore, sanctions will not enter the social welfare function. I discuss the case of partially non-transferable (i.e. costly) sanctions, in section 5.A.

### C. Deterrence

Deterrence, in and of itself, is not an objective. Deterrence serves to reduce the number of victims by making crime an undesirable option for potential offenders. As equations (5) and (6) suggest, the choice of  $s(0)$  and  $s(1)$  will determine the level of deterrence, and accordingly the frequency of crime. Let  $b^*(s(0), s(1))$  denote that individual who is indifferent between abstaining from and committing crime. Then, it trivially follows that individuals with  $b < b^*$  will be deterred from committing crime. Hence,  $b^*$  serves to identify the level of deterrence (i.e.  $F(b^*)$ ).

### D. Objective Function

The pure utilitarian social welfare function, which incorporates victims' psychic costs, offenders' utilities, and deterrence, is simply the sum of all individuals' utilities. Since sanctions ( $s$ ) are transferrable, welfare can be expressed as:

$$\int_{b^*(s(0), s(1))}^{\infty} (b - \alpha g - h + p\mu v(q) + p\mu q r) f(b) db \quad (7)$$

where  $\mu$  is the proportion of detected offenders who apologize.

#### 4. Analysis

The objective function expressed by (7) demonstrates that social welfare depends on the sanction scheme imposed on offenders. In particular, sanctions affect social welfare through three values: (i)  $b^*(s(0), s(1))$  which determines the level of deterrence (i.e.  $F(b^*)$ ), (ii)  $\mu(s(0), s(1))$ , which is the proportion of detected offenders who apologize, and (iii)  $q(s(0), s(1))$  which is the proportion of apologizing individuals who are truly remorseful.

*Ceteris paribus*,  $\mu$  and  $v(q)$  increase social welfare. A high  $\mu$  leads to a greater number of victims receiving apologies, and accordingly lowers victims' psychic costs. The value of each apology ( $v(q)$ ) increases the relief of each victim who receives an apology. Furthermore,  $\mu q$ , which corresponds to the density of individuals who feel remorseful and apologize, increases benefits associated with offenders' relieved conscience.

Next, I determine how  $\mu$  and  $q$  react to sanction schemes. If there is a marginal penalty for apologizing (i.e.  $s(1) > s(0)$ ) then we should expect non-remorseful individuals to abstain from apologizing. Furthermore, as noted in Observation 1 in section 2, there is a threshold marginal penalty (i.e.  $r$ ), such that remorseful individuals apologize only if the price of apologizing is below that threshold. Hence, if the price of apologies are positive but low, only remorseful offenders will apologize.<sup>29</sup> If they are sufficiently high no one will apologize. It should also be clear that if apologizing is rewarded rather than priced, every individual will find it in their best interest to apologize. These are summarized by the following observation:

**Observation 2:** (i) When  $s(1) - s(0) > r$  nobody apologizes, (ii) when  $r \geq s(1) - s(0) > 0$

---

<sup>29</sup>I assume that indifferent individuals apologize. Results would not change if indifferent individuals randomly chose whether or not to apologize. I briefly discuss the significance of assumptions concerning indifference and offer an extension where offenders may have external benefits from apologizing in section 5.D.

only remorseful offenders apologize, and (iii) when  $s(1) - s(0) < 0$  every offender apologizes.

The three cases identified in observation 2 can conveniently be labeled as excessive pricing of apologies ( $s(1) - s(0) > r$ ), pricing of apologies inducing separation ( $r \geq s(1) - s(0) > 0$ ), and rewarding of apologies ( $s(1) - s(0) < 0$ ). An implication of proposition 1 is that these three cases will generate different values of  $q$  and  $\mu$ , which are crucial in determining various effects of apology regimes. Table 1 below, depicts values of  $q$  and  $\mu$  attained under the three apology regimes identified.

**Table 1:**

**Effects of Apology Regimes on  $q$  and  $\mu$**

	$q$	$\mu$
Excessive pricing of apologies	NA	0
Pricing inducing separation	1	$\alpha$
Rewarding of apologies	$\alpha$	1

It should be clear that, *ceteris paribus*, apologies inducing separation lead to higher welfare through their effects on  $q$  and  $\mu$ . To see this, note that excessive pricing of apologies and rewarding of apologies, leads to apologies which provide no information to victims concerning offenders' conscious state (since  $\mu = 0$  in the excessive pricing case and  $q = \alpha$  when apologies are rewarded). On the contrary, apologies inducing separation provide complete information by separating remorseful and non-remorseful offenders (i.e.  $q = 1$ ). Furthermore, apologies inducing separation and rewarding of apologies lead to equal relief on the offenders' side. This follows from the fact that in both regimes, all remorseful offenders apologize and relieve their conscience. However, excessive pricing of apologies prevent all offenders and in particular remorseful offenders from achieving the same relief. Hence, abstracting from the

issue of deterrence, pricing inducing separation should dominate other apology regimes. These observations are summarized by the following proposition.

**Proposition 1:** *For a fixed level of deterrence, apologies inducing separation lead to greater welfare.*

**Proof:** *See Appendix.*

Proposition 1 hints at the conclusion that pricing inducing separation is optimal. But this conclusion cannot yet be reached, because it is not clear whether the *ceteris paribus* condition necessarily holds. That is to say, if excessive pricing or rewarding of apologies lead to gains not associated with  $\mu$  and  $q$ , then it may well be the case that apologies inducing separation are sub-optimal. I have noted earlier that social welfare is affected through a third and final source, namely deterrence ( $F(b^*(s(0), s(1)))$ ). Proposition 2 below, shows that separation does not impact the possibility of achieving any desired level of deterrence. Although a proof of this proposition is provided in the appendix, it is worth noting that this proof does take into account that deterrence levels in rewarding regimes are affected by the knowledge that penalties will be mitigated if the offender apologizes.

**Proposition 2:** *Any level of deterrence that can be achieved under excessive pricing or rewarding of apologies can also be achieved under apologies inducing separation.*

**Proof:** *See Appendix.*

An implication of this proposition along with previous observations is that apologies inducing separation can be set to achieve greater welfare than in any other case. This follows from the fact that separation leads to higher benefits through offenders' and victims' psychic relief, without compromising the gains associated with deterrence. Next, I formalize this observation and identify optimal prices for apologies, which lead to gains described earlier.

**Proposition 3:** (i) *Optimal sanctions result in pricing inducing separation, and (ii) optimal prices for apologies are positive and not greater than  $r$ .*

**Proof:** *See Appendix.*

Proposition 3 summarizes the main result. If remorse is to be relevant in the determination of sanctions, then individuals who assert that they are remorseful should be *punished*<sup>30</sup> more severely. This follows from the presumption that remorseful individuals have more to gain by apologizing than non-remorseful individuals. Hence, slightly increasing the punishment when individuals apologize has the effect of separating remorseful and non-remorseful individuals. This leads to credible apologies. As long as credibility of apologies are valued by victims, it follows that pricing apologies is optimal.

In deriving this result, I abstracted from a number of issues to simplify the analysis. In the next section, I discuss the robustness of results by considering the likely effects of a few more realistic assumptions.

## 5. Discussion

### A. Effects of Costly Sanctions

As I have briefly remarked in the introduction, my main analysis overlooks the fact that sanctions can be costly to impose. Sanctions are certainly costly when they involve imprisonment, and may be costly in other circumstances where the imposition or collection of monetary fines are costly. In these cases, reducing the penalty for any group of individuals will result in savings from costs of sanctioning. In particular, the same will be true for reductions in the penalty for offenders who elect to apologize. However, as I demonstrated in previous sections, there is no good reason to achieve such savings by lowering the penalty for apolo-

---

<sup>30</sup>See, section 5.F. *infra*, discussing how *punishment* for apologizing can take the form of additional imprisonment, civil service, monetary transfers, or a combination thereof.

gizing, as opposed to employing a priced apology regime and lowering the penalty for every offender. Reducing the penalties for apologizing leads to problems I have identified earlier (i.e. non-credible apologies) and lowering deterrence, whereas uniform reductions in penalties only lower deterrence.

As demonstrated in observation 2, rewarding apologies incentivizes every individual to apologize. Hence, gains arising from a credible apology mechanism are lost when there are rewards for apologies. This result is independent of whether or not sanctions are costly. Therefore, a priced apology regime dominates a regime in which apologies are rewarded in terms of gains from a credible apology system (see proposition 1 and preceding text). Furthermore, when sanctions are costly, such costs are a function only of the frequency of crime (alternatively of deterrence). This follows from the fact that the same factors which determine deterrence, also determine costs of sanctions and in the same way.<sup>31</sup> This implies that for any given level of deterrence, costs associated with the imposition of sanctions are equal across regimes. But as demonstrated in proposition 2, any level of deterrence that can be achieved when apologies are

---

<sup>31</sup>To verify my claim that costs of imposing sanctions are dependent only on the level of deterrence, let  $\sigma$  be the per unit cost of imposing a sanction. Accordingly, in any regime, total costs of sanctions are given by:

$$C = \sigma p(1 - F(b^*))[\mu s(1) + (1 - \mu)s(0)]$$

Now let  $b^{*p}$  and  $b^{*r}$  respectively denote the critical  $b$ 's in an apology regime inducing separation and a regime which rewards apologies. Then, the expressions for these values are given by:

$$\begin{aligned} b^{*p} &= \alpha[g - pr] + p[\alpha s^p(1) + (1 - \alpha)s^p(0)] \\ b^{*r} &= \alpha[g - pr] + p[s^r(1)] \end{aligned}$$

where  $s^p$  and  $s^r$  respectively denote sanctions imposed in an apology regime inducing separation and a regime which rewards apologies. It follows that  $b^{*r} = b^{*p}$  if and only if  $s^r(1) = \alpha s^p(1) + (1 - \alpha)s^p(0)$ , and as illustrated in table 1:

$$\mu = \begin{array}{ll} \alpha & \text{for priced apologies} \\ 1 & \text{for rewarding of apologies} \end{array}$$

Accordingly,  $C$  is a function only of  $b^*$ .

rewarded can also be achieved by pricing apologies. Therefore, for any sanction scheme which rewards apologies, one can find a priced apology regime which results in the same amount of costs associated with sanctions, the same amount of deterrence, but higher gains from a credible apology system. Accordingly, my main analysis extends to cases where the imposition of sanctions are costly.

## **B. Self-Reporting and Plea-Bargaining**

In the modelling sections of this paper, I implicitly assumed that self-reporting is not possible. However, results would not change even when offenders may self-report. As explained in previous sections and formalized in proposition 2, pricing apologies does not preclude the possibility of achieving a desired level of deterrence. A similar reasoning would imply that pricing apologies do not prevent policy makers from achieving other goals that can be targeted by self-reporting (for instance reducing detection costs).

An easy way to verify this conjecture is by imagining a legal regime, where, for a given crime, there is a menu of sanctions: One for individuals who self-report, one for individuals who plea-bargain, and one for individuals who go to trial and are convicted. Applying the reasoning in this paper to this legal regime implies that giving offenders the right to apologize and sanctioning them slightly more severely would result in gains due to a credible apology system, without changing the distribution of individuals who choose to self-report, plea-bargain or go to trial. Accordingly, the existence of alternative procedural mechanisms is not likely to influence the results derived earlier.

## **C. The Harm Exceeds Every Offender's Benefit**

An implicit assumption in the modelling section was that some individuals have benefits which exceed the harm associated with crime. However, results should not be affected when

$b < h$  for all individuals. As formalized in propositions 1 and 2, pricing apologies leads to gains for any given level of deterrence and all levels of deterrence that can be achieved under alternative regimes can be obtained in a regime which prices apologies. Hence, results will continue to hold, regardless of the optimal level of deterrence which would be affected by the relationship between  $f(b)$ ,  $h$ , and whether  $p$  is exogenous.

#### **D. External Benefits from Apologizing and Assumptions Concerning Indifference**

One may object to my analysis, because it ignores potential external benefits associated with apologizing. These benefits may include greater chances of being employed in the future or a lower level of stigma. In other words, there may be external benefits which can be received by both remorseful and non-remorseful offenders. My results rely not on the absence of such external benefits, but on the assumption that remorseful individuals have *more* to gain by apologizing than non-remorseful individuals. Hence, the existence of external benefits associated with apologizing which are receivable by all offenders would not affect results. In fact, the existence of such benefits would strengthen results. To elaborate on the last point, I will briefly discuss assumptions concerning indifference and their relevance.

In footnote 29, I elucidated my assumption that indifferent individuals choose to apologize, although results would not change if they randomly chose whether or not to apologize. If, to the contrary, indifferent individuals would never choose to apologize, uniform sanctions as well as higher penalties for apologizing would achieve separation. In sum, my assumption concerning the behavior of indifferent individuals affects the conclusion as to whether uniform sanctions are optimal. This is not very harmful, because priced apologies are optimal under *any* assumption concerning the behavior of indifferent individuals whereas uniform penalties

are optimal only under *a specific* assumption. But when external benefits are present, uniform sanctions do not achieve separation even if indifferent individuals abstain from apologizing.

To verify this conjecture, consider external benefits of  $w > 0$ . In this case, all remorseful individuals will apologize if  $w + g > s(1) - s(0)$ , where the strict inequality reflects the assumption that indifferent individuals do not apologize. On the other hand, non-remorseful individuals will not apologize if  $s(1) - s(0) \geq w$ . Hence, separation will be achieved only if  $w + g > s(1) - s(0) \geq w > 0$ . Hence, uniform sanctions cannot achieve separation, even when indifferent individuals abstain from apologizing.

### E. Social Benefits from Insincere Apologies

In the previous sections I assume that victims have realistic expectations and that they only value sincere apologies. As such, a social welfare function emerges where insincere apologies are not valued. One may legitimately debate the truth value of this assumption and whether it makes sense to use a social welfare function which does not incorporate potential gains from insincere as well as sincere apologies.

Two particular sources of sizeable benefits from apologies (insincere and sincere) comes to mind; (i) *Rehabilitative Benefits*: Insincere apologies may serve to rehabilitate non-remorseful offenders, and (ii) *Palliative Benefits*: "Victims may take pleasure in the humiliation a perpetrator must endure to *humble* himself and apologize. Simply hearing words of apology incanted may have a palliative effect. Whatever the reason, it seems likely that the value of ... an insincere apology is significantly greater than zero."<sup>32</sup> When these are accounted for, a system that rewards apologies has an advantage over a system that prices apologies; it produces more

---

<sup>32</sup>These words belong to an anonymous referee, whom I would like to thank for elegantly describing the potential value from insincere apologies.

apologies and therefore greater rehabilitative and palliative benefits. This being the case, one can no longer immediately conclude that a system that prices apologies dominates one that rewards them. However, the model presented in the main sections of this paper can be easily modified to incorporate such benefits and to determine conditions under which pricing apologies is optimal. This is achieved via the following proposition.

**Proposition 4:** *Let  $i$  denote the sum of rehabilitative and palliative benefits of an apology, then a pricing regime (inducing separation) is optimal iff  $\frac{\alpha}{(1-\alpha)} \geq \frac{i}{v(1)}$ .*

**Proof:** *See Appendix.*

As one would expect, under this new set of assumptions, optimality of pricing regimes depends on the ratio between the value attached to credible apologies and the sum of rehabilitative and palliative benefits. Less expected perhaps, is a second implication of Proposition 4: Pricing regimes are more likely to be optimal if the proportion of remorseful offenders is high. This is perhaps a counter-intuitive result at first glance, because one would expect the problem of sorting the non-remorseful from the truly remorseful offender to become less important as the proportion of remorseful offenders increases. But when pricing regimes are viewed from a different perspective the result becomes less paradoxical. A pricing regime is one that enables the truly remorseful offender to psychologically relieve his victim by making credible apologies possible. As such, the value of such a regime is naturally increasing in the number of truly remorseful offenders.

Whether this optimality condition is satisfied is naturally an empirical question, and one that is hard to answer. Rehabilitative justice and therapeutic jurisprudence scholars may argue that  $i$  is very big and  $v(1)$  is comparatively small, whereas a rational choice theorist may argue the opposite. It is worth pointing out one thing, however. This analysis incorporates potential

benefits associated with insincere apologies (i.e.  $(1 - \alpha)i$ ) without incorporating their potential costs. As stated earlier in the introduction, there are at least two plausible sources of social costs that come with insincere apologies.

First, a regime that rewards apologies runs the danger of generating asymmetric ex-ante incentives to individuals who know that they can fake remorse, and those who cannot. And as is demonstrated in the existing literature, such asymmetries generate costs associated with under- as well as over deterrence.<sup>33</sup> Second, insincere apologies may magnify the anger and hate of victims. I have abstracted from these additional sources of costs associated with rewarding regimes in section 3.A. where I assumed that  $v(\alpha) = 0$ . Accordingly, the size and sign of  $i$  becomes debatable when one considers additional sources of potential negative effects of insincere apologies.

Nevertheless, if one believes that  $i$  is positive, proposition 4 tells us that if (i) a large proportion of offenders are remorseful, and/or (ii) the gains from credible apologies are large relative to the rehabilitative and palliative benefits of apologies, then pricing apologies is superior to regimes that reward apologies.

## **F. Ways to Price Apologies**

The previous sections demonstrate that pricing regimes can generate a credible apology system. This requires making apologies costly to offenders. There are several methods that can be used to generate such costs to offenders. The first, and most straightforward way, is to have sentencing guidelines call for some increase (that is fixed or proportional to the base sanction) in the sanction to be applied to an offender who chooses to apologize. There may be a few problems with this method, especially if one considers informational problems that have

---

<sup>33</sup>See note 6, *supra*.

been assumed away in the model presented.

Offenders may vary substantially from each other in their feelings of remorse. Hence, a uniform pricing system for apologies would presumably be a poor method for signalling offenders' intensity of remorse. Furthermore, since remorseful offenders voluntarily choose to incur additional costs for apologizing, a pre-determined penalty handed down by a social planner does not appear to be necessary for a pricing system to function.

One may consider alternative systems of apologies, such as monetary transfers from the offender to the victim. Such transfers would solve the problem of having a uniform penalty system for offenders who feel varying degrees of remorse, and would allow offenders to reveal how much they regret their actions. This method seems to be susceptible to various criticisms. Most importantly, many scholars may argue that having the offender pay the victim would be a species of *commodification*, the transformation of previously non-commercial concepts into objects that could be purchased, which may cause the apology to have the opposite of the intended effect.<sup>34</sup> The criticism becomes very clear when one considers a sex offender offering his victim a monetary sum to apologize.

Another alternative is to have the offender transfer a sum of his choosing to a charity or organization that supports victims of crimes. This method is presumably not prone to commodification criticisms. A problem with this alternative, and any other method that involves monetary payments, is that it is inaccessible to wealth constrained offenders.

A third alternative is to have the offender select his own sentence enhancement, or choose to engage in a type of civil service. By doing so, an offender can choose an enhancement

---

<sup>34</sup>See, e.g., Ertman and Williams (2005) referring to this concept, and bringing together several pieces that criticize commodification processes, and others which defend the creation of previously non-existent markets.

(or civil service) that would inform his victim of how remorseful he really is. Moreover, this method does not suffer from commodification criticisms. It does suffer from another criticism, however. Monetary transfers, as is well known in the law enforcement literature, are less costly than non-monetary sanctions, because they are transferable. One person's loss is another person's gain.

A hybrid system, and one that appears to be superior to all other systems discussed above, is one where the offender can elect to (i) make donations to a charity, (ii) choose a sentence enhancement or commit to civil service, or (iii) a combination of both. Such a system would eliminate the problem of inaccessibility by wealth constrained offenders and would partially (in cases where the offender chooses (i)) eliminate costs associated with non-transferability.

## **6. Conclusion**

I analyzed optimal punishment schemes in a stylized model where offenders have the option of expressing remorse. I concluded that, contrary to current legal practice, it may be optimal to price apologies rather than rewarding them. I relied on the presumption that remorseful offenders have more to gain than non-remorseful individuals by apologizing (i.e. a relieved conscience). When this is true, it is possible to separate remorseful and non-remorseful offenders by pricing apologies, which leads to credible apologies. When it is assumed that insincere apologies carry no social value, it follows that such separation is optimal, since it creates valuable and credible apologies and no social costs. Varying intensities of remorse were assumed away in the formal model to simplify the analysis, and the importance of this assumption was discussed separately and informally. I then discussed robustness of results, and argued that my conclusions extend to broader sets of assumptions.

My formal analysis was purely utilitarian, but it incorporated agents who are not narrowly

self-serving. In particular, I incorporated potential offenders who may feel remorse, and also care about apologizing. Hence, my results suggest that the inconsistency between current legal practices concerning manifestations of remorse and optimal punishments identified by economic models is not the product of assuming narrowly self-serving individuals. Therefore, if one seeks justifications for current legal practices concerning remorse, one should probably search these in benefits associated with insincere apologies.

I labeled two particular sources of potentially sizeable benefits that may arise from insincere apologies as rehabilitative and palliative benefits. Once these are accounted for, the optimality of pricing regimes depends on a trade-off between such benefits and the value of having a credible apology system. A second factor that determines whether a pricing regime is optimal is the proportion of offenders who feel remorse. This is because any positive social value attached to insincere apologies and the value of a credible apology system are both increasing in the number of individuals capable of making such apologies (i.e. non-remorseful and remorseful offenders respectively). I also pointed out that the size and sign of the effects of insincere apologies is debatable, because they may generate social costs associated with under- and over-deterrence, as well as additional psychic costs for the victims.

In sum, current legal practices concerning remorse and apologies may be sub-optimal from a utilitarian perspective, and it may be better to price rather than reward apologies. This conclusion is not the product of assuming narrowly self-serving individuals, and follows whenever credibility benefits outweigh potential palliative and rehabilitative benefits of insincere apologies.

## **Appendix**

**Proof of Proposition 1:** *Plugging in the values for  $q$  and  $\mu$  as described in Table 1 into the objective function expressed in (7) reveals that social welfare as a function of how apologies are treated is given by:*

$$\begin{aligned}
& \int_{b^*(s(0),s(1))}^{\infty} (b - \alpha g - h) f(b) db && \text{under Excessive pricing of apologies} \\
& \int_{b^*(s(0),s(1))}^{\infty} (b - \alpha g - h + p\alpha v(1) + p\alpha r) f(b) db && \text{under Pricing inducing separation} \\
& \int_{b^*(s(0),s(1))}^{\infty} (b - \alpha g - h + p\alpha r) f(b) db && \text{under Rewarding of apologies}
\end{aligned} \tag{A.1}$$

*Fixing the level of deterrence is equivalent to assuming  $b^*$  is equal across regimes, in which case it trivially follows that Pricing inducing separation leads to greater welfare. Q.E.D.*

**Proof of Proposition 2:** *Equations (5) and (6) imply that  $b^*$ , the value which determines deterrence, as a function of how apologies are treated is given by:*

$$b^* = \begin{cases} b - \alpha g - ps^e(0) & \text{under Excessive pricing of apologies} \\ b - \alpha g - p(\alpha s^p(1) + (1 - \alpha)s^p(0)) + p\alpha r & \text{under Pricing inducing separation} \\ b - \alpha g - ps^r(1) + p\alpha r & \text{under Rewarding of apologies} \end{cases} \tag{A.2}$$

*where  $s^e, s^p$  and  $s^r$  respectively denote sanctions under excessive pricing, separation and rewarding regimes.*

*For any  $s^e(0)$ , there is a pair of  $s^p(1), s^p(0)$  such that  $\alpha s^p(1) + (1 - \alpha)s^p(0) + \alpha r = s^e(0)$  satisfying  $r \geq s^p(1) - s^p(0) > 0$ . Similarly, for any  $s^r(1)$  there is a pair of  $s^p(1), s^p(0)$  such that  $\alpha s^p(1) + (1 - \alpha)s^p(0) = s^r(1)$  satisfying  $r \geq s^p(1) - s^p(0) > 0$ . Hence, any level of deterrence that can be achieved under excessive pricing or rewarding of apologies can also be achieved under apologies inducing separation. Q.E.D.*

**Proof of Proposition 3:** (i) Choose any sanction pair  $z$  which generates a non-separation regime. Proposition 2 implies that there is a corresponding sanction pair  $s$  which generates a separation regime and the same level of deterrence. Proposition 1 implies that this  $s$  leads to greater welfare than  $z$ .

(ii) Proposition 1 implies that separation requires  $r \geq s(1) - s(0) > 0$ . Part (i) of this proposition states that optimality requires separation, hence the optimal price for apologies (i.e.  $s(1) - s(0)$ ) must satisfy  $r \geq s(1) - s(0) > 0$ . *Q.E.D.*

**Proof of Proposition 4:** First note that for purposes of determining optimality, excessive pricing regimes can be ignored due to results obtained in propositions 1-3. Next, given  $b^*$  (the value describing deterrence) (A.7.) can be modified to express social welfare under separation and rewarding regimes.

$$\begin{aligned} \int_{b^*(s(0),s(1))}^{\infty} (b - \alpha g - h + p\alpha v(1) + p\alpha r + p\alpha i) f(b) db & \quad \text{under Pricing inducing separation} \\ \int_{b^*(s(0),s(1))}^{\infty} (b - \alpha g - h + p\alpha r + p i) f(b) db & \quad \text{under Rewarding of apologies} \end{aligned} \tag{A.3.}$$

As such pricing regimes achieve higher welfare than rewarding regimes iff  $\frac{\alpha}{(1-\alpha)} \geq \frac{i}{v(1)}$ . Furthermore, as proven in proposition 2 any  $b^*$  can be achieved by both regimes. Hence, a pricing regime is optimal iff  $\frac{\alpha}{(1-\alpha)} \geq \frac{i}{v(1)}$ . *Q.E.D.*

## References

- [1] Adler, Jacob. 1992. The urgings of conscience: a theory of punishment. Philadelphia: Temple University Press.

- [2] Becker, G. S. 1968. "Crime and Punishment: An Economic Approach," 76 *Journal of Political Economy*, 169-217.
- [3] Bernick, Michael S. 1982. "Of Crime and Conscience," 68 *A.B.A. Journal*, 306-309.
- [4] Bibas, Stephanos and Richard A. Bierschbach. 2004. "Integrating Remorse and Apology into Criminal Procedure," 114 *The Yale Law Journal*, 85-148.
- [5] Cochran, Robert F. Jr. 1998. "Crime, Confession, and the Counselor-At-Law: Lessons from Dostoyevsky," 35 *Houston Law Review*, 327-397.
- [6] Ertman, Martha M., and Joan C. Williams. 2005. "Rethinking commodification: cases and readings in law and culture." New York: New York University Press.
- [7] Frank, Robert H. "Departures from Rational Choice: With and Without Regret" in Parisi, F., and Smith, V. L. (2005). *The law and economics of irrational behavior*. Stanford, Calif: Stanford Economics and Finance, Stanford University Press.
- [8] Garoupa, Nuno. 1997. "The Theory of Optimal Law Enforcement," 11 *Journal of Economic Surveys*, 267-95.
- [9] Innes, Robert. 1999. "Remediation and self-reporting in optimal law enforcement," 72 *Journal of Public Economics*, 379-393.
- [10] Innes, Robert. 2000. "Self-Reporting in Optimal Law Enforcement When Violators Have Heterogeneous Probabilities of Apprehension," 29 *The Journal of Legal Studies*, 287-300.
- [11] Innes, Robert. 2001. "Violator Avoidance Activities and Self-Reporting in Optimal Law Enforcement." 17 *Journal of Law, Economics and Organization*, 239-56.

- [12] Kaplow, Louis. 1990. "A Note on the Optimal Use of Nonmonetary Sanctions," 42 *Journal of Public Economics*, 245-247.
- [13] Kaplow, Louis and Steven Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior," 102 *Journal of Political Economy*, 583-606.
- [14] Mungan, Murat C. 2010. "A Note on the Effects of State-Dependent Benefits on Optimal Law Enforcement," 6 *Review of Law and Economics*, 97-106.
- [15] Mungan, Murat C. 2011. "Welfare Enhancing Regulation Exemptions" *International Review of Law and Economics*, In Press, DOI:10.1016/j.irle.2011.06.002.
- [16] Netzig, Lutz and Thomas Trenczek. "Restorative Justice as Participation: Theory, Law, Experience and Research," in Galaway, B., & Hudson, J. (1996). *Restorative justice: International perspectives*. Monsey, NY: Criminal Justice Press.
- [17] O'Hear, Michael M. Remorse, 1997. "Remorse, Cooperation, and "Acceptance of Responsibility": The Structure, Implementation, and Reform of Section 3E1.1 of the Federal Sentencing Guidelines," 91 *Northwestern University Law Review*, 1507-1573.
- [18] Pipes, Randolph B. and Marci Alessi. 1999. "Remorse and a Previously Punished Offense in Assignment of Punishment and Estimated Likelihood of a Repeated Offense," 85 *Psychological Reports* 246-248.
- [19] Polinsky, A. Mitchell and Steven Shavell. 1984. "The Optimal Use of Fines and Imprisonment," 24 *Journal of Public Economics*, 89-99.
- [20] Polinsky, A. Mitchell and Steven Shavell. 2000. "The Economic Theory of Public Enforcement of Law," 38 *Journal of Economic Literature*, 45-76.

- [21] Raskolnikov, Alex. 2009. "Revealing choices: Using taxpayer choice to target tax enforcement," 109 *Columbia Law Review* 689-754
- [22] Sarat, Austin. "Remorse, Responsibility, and Criminal Punishment: An Analysis of Popular Culture" in Bandes, Susan A. 1999. *The passions of law*. Critical America. New York: New York University Press.
- [23] Shavell, Steven. 1987. "The Optimal Use of Nonmonetary Sanctions as a Deterrent," 77 *The American Economic Review* 584-592.
- [24] Shavell, Steven. 1990. "Deterrence and the Punishment of Attempts," 19 *The Journal of Legal Studies*, 435-466.
- [25] Simons, Michael A. 2003. "Retribution for Rats: Cooperation, Punishment, and Atonement," 56 *Vanderbilt Law Review*, 1-54.
- [26] Stephenson, Matthew. 2006. "A Costly Signaling Theory of "Hard Look" Judicial Review," 58 *Administrative Law Review* 753-814.
- [27] Strang, Heather and Lawrence W. Sherman. 2003. "Repairing the Harm: Victims and Restorative Justice" 15 *Utah Law Review* 15-42.
- [28] Tasioulas, John. 2007. "Repentance and the Liberal State", 4 *Ohio State Journal of Criminal Law*, 487-521.
- [29] Tieger, Alan. 2003. "Remorse and Mitigation in the International Criminal Tribunal for the Former Yugoslavia," 16 *Leiden Journal of International Law* 777-786.

- [30] Toffel, Michael W. and Jodi L. Short. 2011. "Coming Clean and Cleaning Up: Does Voluntary Self-Reporting Indicate Effective Self-Policing?", Georgetown Public Law and Legal Theory Research Paper No. 11-1
- [31] Ward, Bryan H. 2006. "Sentencing Without Remorse," 38 Loyola University Chicago Law Journal, 131-167.