The Sorry Clause*

By Vatsalya Srivastava

Tilburg University†

21 April, 2015

Abstract

When players face uncertainty in choosing actions, undesirable outcomes cannot be avoided. Accidental defections caused by uncertainty that does not depend on the level of care, require a mechanism to reconcile the players. This paper posits the existence of a 'sorry equilibrium' that relies on a costly apology for self-identification of accidental defection in a social dilemma. The cost of such an apology is calculated and the outcomes of such an equilibrium are compared to those from other bilateral social governance mechanisms and formal legal systems. It is argued that with the possibility of accidental defections other social mechanisms are inadequate, while formal legal systems can generate perverse incentives.

JEL classification: D08, K04, Z01

Keywords: Apology, Sorry, Social Norms, Uncertainty, Social Welfare, Governance systems, Legal Institutions

---

* I am grateful to Dr. Jens Prufer for his insightful comments and feedback. All errors are my own.

† CentER, TILEC, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands; v.srivastava@uvt.nl.

1 Introduction

It stands to reason that an efficient legal system is critical to the health of any modern economy. Courts empowered by the law of the land defend property rights, enforce contracts and limit rent-seeking; creating the requisite conditions for productive economic activity. However, they can also inflict onerous costs on the economy and society, creating impediments to the existence of a prosperous society. In this regard some figures from the legal systems of United States of America and India are illustrative.

The criminal justice system in the U.S has 6.94 million people or 1 in every 35 American adults under its purview, of which approximately 2 million are incarcerated (BJS 2012). A report by VERA Institute of Justice[1] reports that the cost of maintaining prisons in 40 states that participated in their survey was $39 Billion in 2010. These statistics are staggering, but the socio-political impact of high incarceration rates and the huge burden they put on the taxpayer notwithstanding, these number do not reflect the full scale and scope of the impact that the legal system has on the economy at large.

The U.S. Chamber Institute for Legal Reform (ILR)[2] estimates that lawsuits cost the U.S economy $264 Billion per year, implying a cost of $850 per capita. The 9th Fulbright Annual Litigation Trends Survey[3] found that in the 12 months leading up to their 2012 survey, 32% of U.S. companies had faced more than 20 suits. Further, these lawsuits, as the ILR mentions, disproportionately affect small businesses and cost them over $100 billion per year. In a survey of small business owners they find that 1 in 3 report that they have been sued or threatened with a lawsuit – and if sued, more than two-thirds say they would likely have to pass legal costs on to consumers, reduce employee benefits and hold off on hiring new employees.
The Pacific Research Institute (PRI)[4] estimates the static costs of litigation alone; which include annual damage awards, plaintiff attorneys' fees, defense costs, administrative costs and deadweight costs; to be as high as $328 Billion per year. The report also estimates dynamic costs of litigation; which include impact on R&D spending, the costs of defensive medicine and the related rise in health-care spending and reduced access to health care, and the loss of output from deaths due to excess liability. The adverse

[1] The Price of Prisons: What incarceration costs taxpayers (Henrichson and Ruth 2012)
[2] http://www.instituteforlegalreform.com/issues/lawsuit-abuse-impact
[3] http://www.nortonrosefulbright.com/news/93066/fulbrights-9th-annual-litigation-trends-survey-litigation-bounces-back-regulation-hits-high-us-release
[4] Jackpot Justice: The true cost of America's tort system (McQuillan et al. 2007)

impact on innovations alone is estimated to cost $367 Billion in lost sales of new products every year, concentrated primarily in 13 industries where tort costs have exceeded their optimal levels (Viscusi and Moore 1993). The PRI estimates the combined static and dynamic costs of the American tort system at a humungous $865 Billion per year.

The Indian Judicial System, servicing the world's second most populous country is another massive legal institution. There are about 20 million pending cases in lower courts and 3.2 million such cases in the high courts (Hazra and Micevska 2004), more recent estimates put the total number closer to 30 million (Times of India, 2013)[5]. 70% of the incarcerated population is under-trial and many for those awaiting trials have been in jail longer than a formal sentence would have required them to be (Krishnan and Raj Kumar 2010). The costs of such a logjam in the courts, both social and economic, are so large that while often spoken of, hardly any comprehensive and reliable estimates are available.

In India, the limited capacity of the legal system is a binding constraint on the number of cases that can be handled by the courts. In the US, a relatively more efficient legal system with a propensity to grant multi-million dollar damages and commit people to supervision has probably created perverse incentives whereby, it is individually rational for citizens to access the courts even when it is not socially optimal[6] to do so. The nature of the problems facing the legal systems of the two countries and their sources are disparate, but it is apparent that both the systems require improvements.

In contrast to the U.S and India, in the Japanese legal system the "attitude of the accused – willingness to apologize and confess – is crucial to the decision of whether to prosecute or not" (Haley 1982). In fact, in Japan at one time "approximately 33% of all cases non-traffic offences" were suspended by the prosecution (Haley 1982; Stephens 2008). Further, in the cases that are prosecuted and the defendant is found guilty, the court suspends the jail sentences in more than two-thirds of such cases. Apology is therefore used as an informal sanction which reduces the likelihood that a dispute will be taken to court (Haley 1998).

[5] In 2010, a high court judge had remarked that the courts would take another 320 years to clear the backlog (TOI 2010)

[6] The optimal here refers to social optimal where (marginal cost)=(marginal benefit) given the law and institutional capacity. This is not to deny the role of supply side activities in altering the optimal level.

The Japanese case is illustrative of the fact that formal legal institutions are not indispensible to covenants of cooperation that allow for productive economic activities (Ostrom, Walker, and Gardner 1992; Ostrom 2000; Posner 1997). Collective action, governed by s*ocial norms* and fostered by internalized pro-social preferences offer an alternative to the third party governance mechanisms of courts where outsiders[7] rely on the threat of punishment to enforce good conduct (Dixit 2009). Such governance mechanisms rely on the widespread cognizance of pro-social norms that allow for enforcement. These norms reflect a social contract, an implicit agreement among the members of a society. One such instance of a ubiquitous social contract is that of an apology (sorry)[8,9]. In its various expressions in different languages it is immediately recognized to be an expression of regretful acknowledgement of offence or failure by both the party offering it and the party receiving it. But, widespread acceptance is not the only criterion that lends credence to apology being a non-trivial social norm that governs actions. Its use in the Japanese legal system provides a great example for apologies being effective in a wide-variety of scenarios; resolving disputes and fostering less acrimonious attitudes. But lest the case of Japan be considered an isolated cultural anomaly, there are other examples from all over the world.

In the U.S., evidence from a study on settlement offers at the time of an accident indicates the relevance of apologies. 73% of the respondents would accept the settlement offered if a full apology was tendered compared to 52% when no apology was offered (Robbennolt 2003). Since 1989, community conferences in New Zealand have dealt with more than half of the juvenile offenders in an alternative system that relies on getting the victim and the offender to meet and utilizes apologies for dispute resolution and victim rehabilitation (Scheff 1998). In Australia the same system is to put to use for both adult and juvenile crimes. In fact, it is claimed that tort plaintiffs often profess that what they really wanted was an apology and brought suit only when it was not forthcoming (Shuman 2000)[10].

The role of apologies has been significant enough to spawn a host of literature in different disciplines. There is evidence of the importance of apologies from business

[7] Outsiders are people who are not directly party to a transaction. Such transaction could be willing (sale, purchase), unwilling (theft, robbery), between known people or strangers. But it must involve more than 2 people.
[8] The term apology is considered to be synonymous with sorry and the terms are used interchangeably in the rest of the paper.
[9] For an extensive overview of the role of social norms (other than apology) in economic governance refer (Dixit 2009).
[10] The source of the claim concedes that such remarks by plaintiffs could at least in part be public posturing.

(Folkes 1984; Hearit 1994), social and child psychology (Darby, Bruce W. 1982; Schlenker and Darby 1981; Schlenker 1980) linguistics (Cohen, A. D., Olshtain 1985; Reiter 2000), history (Trouillot 2000; Gibney 2008) and politics (Lind 2011; Nytagodien and Neal 2004). Legal scholars have evaluated and extolled the virtues of apologies in keeping down crime rates in general and recidivism rates in particular (Haley 1998); and providing cathartic relief to affected parties (Shuman 2000; Keeva 1999). There is also evidence from experiments conducted by economists (Fischbacher and Utikal 2013; Ho 2012) that apologies are a useful conflict resolution mechanism.

This plethora of evidence supporting the role of apology in governing interaction and maintaining social organization, establishes a strong case for an economic analysis of the phenomena. This paper is an attempt to conduct such an analysis, by developing a game-theoretic framework to rationalize the existence, usage and effectiveness of apologies. This framework will allow for addressing the many questions that stand in need of investigation.

What in the nature of socio-economic interactions necessitates the need for apology? What is the cost of such an apology? Is apology an equilibrium outcome? If so, what parameter configurations allow for the existence of such equilibrium? How do the welfare outcomes in a sorry equilibrium compare to that in other modes of community governance, like ostracizing defaulting members? How do the welfare outcomes compare to that from formal legal institutions like courts? What additional constraints do courts face compared to apologies?

The next section will review the relevant extant literature on community interaction and governance while identifying the unique characteristics of apologies. Section 3 outlines the model; lays out the assumptions and the proposed equilibrium. Section 4 analyzes the model and discusses results. Section 5 compares the welfare outcomes of multiple community enforcement mechanisms. Section 6 outlines a model for court governance and compares the outcomes to those from the sorry equilibrium. Section 7 concludes with a discussion on the limitations of the results posited.

2 Characterizing an Apology

The existence of pro-social preferences, norms and related enforcement mechanisms can allow for bi-lateral or multi-lateral interactions to overcome the problems of prisoner's dilemmas[11] (Dixit 2009; Greif 1993; Bowles and Gintis 1998). The mechanism that allows for a system of cooperation has been modeled in the literature as self-reinforcing equilibriums that are made possible by repeated interactions between the agents. Some of these models rely on the possibilities afforded by the *folk theorem* in infinitely repeated interactions (Sobel 2006; Greif 1993; Abreu 1988; D Fudenberg and Maskin 1986). The basic idea is that the stream of payoffs of each player (in a static game) in a cooperative equilibrium exceeds the discounted value of the stream of payoffs from cheating. In infinitely repeated games cheating is dissuaded by some variant of a grim-trigger strategy/ostracism/fines that threatens cheating with a punishment that offsets the one-time gain from cheating. In multi-lateral setting, the reputation[12] of the players also plays an important role. So: if A cheated B; C when interacting with A is obligated to enforce the punishment, such that the threat of such a punishment dissuades A from cheating against B. In situations where reputations are important, case specific ex-post arrangement (Bull 1987) in finitely repeated interactions can also lead to cooperative outcomes. In all such models of cooperation without external enforcement, the punishment never gets played out in the equilibrium, the threat exists merely to ensure that players cooperate in equilibrium. This is also true for the role of legal sanctions in law and economics literature, where sanctions never get played on the equilibrium path, but are put in place to deter cheating (Polinsky and Shavell 1984).

The equilibrium outcomes of such models, however, are at odds with observations of the day-to-day usage of apologies. They are used rather frequently, both in long-standing relationships as well as in interactions with comparative strangers. Apologies are understood to be an acknowledgement of failure or offence but as these theoretical models do not account for the possibility of an offence[13] in equilibrium, it makes operationalizing the concept impossible. This leads to the first 2 unique characteristics of an apology that distinguish it from other forms of social enforcement:

1. The frequent usage of apologies implies that it must be an equilibrium outcome and is not to be relegated to the off-equilibrium path to act as deterrence.
2. In order for an equilibrium to allow for apologies there must be scope for failures/offences in equilibrium. Moreover given the usage of apologies in a

[11] Standard terminology of prisoner's dilemma games offers each player the choice to either cooperate or cheat/defect.
[12] Reputation here refers to the history of a players interactions
[13] Offence here is used to mean cheat/defect

wide-variety of circumstances, the cause of such failures must be sufficiently general to find manifestation in many different types of events.

There exists a literature in experimental economics (Denant-Boemont, Masclet, and Noussair 2007) that shows that in social-dilemmas, given the opportunity, sanctions[14] do get played out, thereby implying that offences and sanctions that accompany them may not be restricted to the off-equilibrium path. However, these results do not offer insights into how such repeated interactions should be modeled to generate sanctions in equilibrium[15]. Further, sanctions by players are premised on other player(s) not conforming to some expected standard and if allowed for in the experimental setting, sanctions lead to avenging-sanctions. Apologies, however, in most social conventions can only be offered, not taken. This implies that an apology is a self-imposed sanction.

3. A village elder can order it or friends can goad for it, but eventually an apology has to be offered by the offender to the offended. This distinguishes it from more conventional punishment that can be met out by some authority or an outside agent. In fact, an apology has a redeeming quality built into it. This characteristic is most visible in social dictums regarding apology that emphasize the need to feel sorry, when an apology is offered.

The recent literature on repeated games with private monitoring and communication (Bhaskar and Obara 2002; Compte 2014) has also generated some insight on the role of apologies. The role of apologies emerges in such games from the lack of observability of actions and randomness of the signals (of actions taken by other players) observed by the players (Ashkenazi-Golan 2004). Deviation by a player cannot always be observed and apologies or confessions serve as costless signals that reduce the severity of punishment in case of detection. The assumption of apologies being costless, however, is problematic. Relevant and useful cheap talk requires that signalers and receivers share their interests to a substantial degree (akin to a coordination game) and social dilemmas do not afford this luxury (Ohtsubo and Watanabe 2009; Ho 2012).

4. Apologizing is costly. This observation while not self-evident is a critical one. The very fact that apologies are rather common implies that they perform the purpose of placating the affected party. Now, if there was no cost attached to

---

[14] Actions that reduce the payoff of other players

[15] There is literature on modeling conditional cooperation, inequity aversion and other behavioral aspects of players to reflect some similar results. However, discussion of these suggestions is not considered relevant to the argument being developed here.

apologizing, the party causing the harm can always roll off an apology and consequently there is no reason why the affected party should take any apology seriously. So there must be some psychological, social or even monetary costs attached to apologizing. This argument is akin to claiming, that in general, apologizing is not cheap talk.

5. There are different levels of sorry. The social psychology literature differentiates between at least 5 different levels of apologies (Schlenker and Darby 1981). The easiest way to think of this claim is to imagine the varying levels of sincerity with which sorry can be said. A simple sorry, or a profuse apology or even buying a sorry (greeting) card to convey the apology. This implies that the degree (cost incurred) of sorry can change to reflect some underlying differences in the circumstances.

Further, such a setting does not allow for what would be considered a genuine case for an apology where an unintentional mistake is committed. This is because is this class of games the need for an apology is motivated by very specific situations where neither players knows which actions have been played with certainty. An alternative is provided by studies in social psychology that posit that an apology is a plea for the offended to not take the event as representative of the intentions and character of the offender (Darby, Bruce W. 1982; Schlenker 1980; Schlenker and Darby 1981).

6. An apology conveys an acknowledgement of harm or offence caused, but it also is an attempt to establish a lack of intent on part of the offending party. So, while the harm caused is not disputed, the contention is that any such harm was accidental and not the result of an intentional attack. This '*separation of intent and outcome'* is a significant departure from most economic models of social interaction.

This separation of intent from outcome is probably the most significant motivation for the existence of apologies. In game theoretic terms such a setting implies that players cannot choose their action[16] with absolute certainty; so while they might intend to cooperate, they end up cheating the other player. There could be multiple reasons for such a separation. Some cases might be induced due to lack of ability; a not so good football player may shoot into a passerby. Some others might be due to lack of information; a new cook might cause a microwave to malfunction. But the most common

[16] Separation of outcome and intent might also result if a player cannot control the payoff of the other player despite being able to control his action.

cause is likely to be the stochasticity innate in many activities. A six-sigma certified firm still produces some lemons and even great athletes despite their best intentions don't always meet expectations. There are cognitive limitations; memory of even the sharpest people fails them once in a while. There are unexpected occurrences; a security guard might have to visit the toilet and the in the process not be able to prevent a theft. There might be unintended consequences; the best new IT network designed for higher speed might crash as one of the millions of variables was not taken into account. All these possibilities can, in general be categorized as accidents of some kind or another.

The possibility of mistakes that such a separation implies has already been used in the literature to explain the need for apologies (Fischbacher and Utikal 2013). The experimental results that emerge from the study are encouraging and that show that "An apology is a strong and cheap device to restore social or economic relationships that have been disturbed" (Fischbacher and Utikal 2013). Players use apologies if their intentions cannot be easily inferred and consequently they fear punishment. However, this particular attempt at explaining the usage of apologies relies on players having other regarding preferences to operationalize it. Such an approach requires additional assumptions about preferences and limits the possibility of exploring why such preferences might have been socialized. It also entails the need for additional parameters, which limits the extent to which the relationship between the cost of apology and gravity of the mistake can be explored.

There is another class of games, with imperfect public information, that model such a separation between intent and outcome without additional behavioral assumption, These games have been explored to provide optimal policy to deal with offences caused accidentally (Rubinstein 1979). Additionally, it has also been shown that a version of the Nash-threat folk theorem exists for such games (Drew Fudenberg, David, and Maskin 1994). However, the strategies suggested in these approaches to resolve the problem of accidental offences are fundamentally different from that of an apology and it is useful to distinguish between them. In any such bilateral game, there is an offending player and an offended player. In the folk-theorem approach, the offended player is responsible for statistically distinguishing between an accidental offence and a deliberate one. While an apology, allows the offending player to take on the burden of resolution.

The legal literature on torts and liability (Brown 1973) also deals with the case of probabilistic outcomes and proposes optimal sanctions for accidents. However, the underlying assumption in such analysis is that the probability of an accident is the function of a choice variable for the player: care[17]. The legal sanctions in this literature are geared towards enforcing the socially optimal level of due care. However, while it is entirely reasonable that the level of care determines some (probably a large) degree of an activity's stochastic nature, in many cases there is some degree of *residual stochasticity*[18] that cannot be controlled. Residual stochasticity in activities requires mechanisms, robust to probabilistic outcomes that resolve undesirable outcomes when they do transpire[19].

3 The Model

The setting for this model is a series of infinitely repeated bilateral games, where the players select their actions simultaneously[20] at every repetition. The stage game for this bilateral interaction is taken to be a simple prisoner's dilemma with the following actions and payoffs:

Table 1: Payoff Matrix of the Prisoner's Dilemma game

1/2	Cooperate (C)	Defect (D)
Cooperate (C)	h, h	l, w
Defect (D)	w, l	d, d

Where: $w > h > 0 > d > l$ and $2h > w + l$

The time between two repetitions is taken to be discrete $t \in \{0, 1, 2 \dots \infty\}$. The rate of time preference for each player is assumed to be constant $\delta \in (0,1)$.

An apology is defined as a self-inflicted cost, undertaken by the offender to convince the offended player that any infractions committed were not intentional. The cost of an apology is captured by the derived parameter s, whose value will depend on other parameters of the model. Further, the cost of the apology is borne by the party apologizing and these costs do not accrue, distinguishing apologies from compensation mechanisms.

This basic model can be adapted to account for the uncertainty resulting from residual stochasticity in the process of playing an action. The inclusion uncertainty also warrants a more precise definition of intention, choosing and playing. In this case, the choice of a player is driven by his intention. So if a player intends to cooperate, he chooses C. But,

[17] Usually represented as p(x), where x is the level of care.
[18] p, not as a function of x
[19] As extent of care does not affect the probability of accidents there is no cost involved in avoiding them and therefore considerations of damages must take primacy.
[20] The game can be extended to depict multi-lateral community interaction, wherein a player might not be playing the same player in every repetition of the game, but all potential players in the community know of the outcome of previous (with some probability) iterations of the stage game.

the uncertainty in action choice ensures that the player cannot ensure that C is actually played out. In effect, the distinction between intention and outcome emphasized earlier is captured by the difference between chosen action and the action that gets played. This also has consequences for the information available to the other player(s). The other player(s) can only observe the action that gets played out, but have no information about the intentions behind it.

The assumptions to incorporate the role of uncertainty in the model are as follows:

1. Player 2 can play the action that he intends to play. For instance, if player 2 intends to Cooperate, he can play Cooperate with certainty.
2. Player 1 cannot ensure with certainty which action will play out. This uncertainty is captured by the parameter p. So, if player 1 chooses to play C, C gets played with probability p and D gets played with probability (1- p). Alternately, if player 1 chooses to play D, D gets played with probability p and C gets played with probability (1- p).
3. $1 \geq p > 0.5$. The laws of probability mandate the first part of the inequality. The second part of the inequality reflect that an action chosen by player 1 is more likely to be played than not.
4. Player 2 can observe the action that gets played by player 1, not the action that player 1 chooses to play.

These assumptions can be incorporated in the stage game to reflect expected payoffs in every period t:

Table 2: Expected Payoffs in the Stage Game

1/2	Cooperate	Defect
Cooperate	$ph+(1-p)w$, $ph+(1-p)l$	$pl+(1-p)d$, $pw+(1-p)d$
Defect	$pw+(1-p)h$, $pl+(1-p)h$	$pd+(1-p)l$, $pd+(1-p)w$

Notation: E^1_{CD} is the expected payoff of the player 1, when he chooses to play C and player 2 plays D:

$E^1_{CD} = pl + (1-p)d$

This is because, while player 2 plays D, player 1 plays C with probability p and D with probability (1-p). The actions of player 1 listed in the matrix are the actions that he chooses to play. So, E^1_{CD} is reflected in the first row, second column of the payoffs. The first row, first column reflect E^1_{CC} and E^2_{CC}.

5. The game is interesting only if $E^2_{CC} > E^2_{DD}$, so it is assumed that $p > \frac{w-l}{h+w-d-l}$[21].

[21] $0.5 < \frac{w-l}{h+w-d-l} < 1$, as $0 < h-d < w-l$.

Timing with-in the stage game in every period t is as follows:

Stage 1: The prisoner's dilemma is played out and payoffs are realized

Stage 2: Player 1 decides whether to apologize or not.

The equilibrium concept used in the analysis that follows is Sub-Game Perfect Nash Equilibrium (SPE) in pure strategies.

3.1 Proposed Equilibrium Strategy with Apology

Player 2 has only one decision to make in every period t, the proposed equilibrium strategy for the repeated game for player 2 is to play C at $t = 0$ and $\forall t \geq 1$:

- Play C if (C, C) is played in (t-1) or player 1 has apologized, else play D

The proposed equilibrium strategy for the repeated game for player 1 is to choose to play C at $t = 0$ and $\forall t \geq 1$:

- If (C, C) is played (t-1), continue to choose to play C
- Else if (D, C) is played out in the previous repetition, apologize by incurring cost s and continue to play C
- Else choose to play D

This strategy[22] is a combination of a form of sanctions (self-inflicted cost of apology) and a trigger strategy. The strategy will constitute an equilibrium if each player is better off choosing this strategy than defecting from it. The required conditions must mandate that each player has an incentive to apologize when a bad outcome transpires and a disincentive to intentionally choose to defect.

4 The Sorry Equilibrium

To show that the proposed strategy constitutes an equilibrium of the game requires the application of backward induction on the decisions made by each player in each period. This requires starting the analysis at stage 2 to determine if the proposed strategy, of apologizing when an accidental defection happens, is incentive compatible for player 1. Further, given the strategy of player 1 at stage 2, it must be incentive compatible for both the players to choose to play C at stage 1 for the proposed strategy to constitute an equilibrium.

4.1 Incentive Compatibility Constraints for Player 1

The assumption of un-observability of intentions creates the need for determining the cost of apology at the margin. This is because an apology can be offered even if the

[22] This paper does not intend to show the uniqueness of this proposed strategy. There can be other strategies involving an apology that may constitute equilibrium. For example, a strategy that allows player 1 to apologize only once for every 2 mistakes and consequently requires that player 2 continue to cooperate after the first mistake knowing that if a second mistake occurs player 1 will apologize, might also constitute an equilibrium. The strategy proposed here is the simplest possible strategy that involves the use of apologies.

defection is intentional: player 1 can choose to defect and if the outcome is (D,C), apologize to claim that the outcome was not intentional. This possibility creates the need to determine the appropriate cost of the apology for a given level of uncertainty (as captured by p) and value of the payoffs. The costs should be mitigated by the possibility that a bad outcome might be unintentional, but high enough to deter intentional deviation from the preferred outcome. This line of argument, is very similar to the one made in Greif (1993), albeit in a different setting,

Stage 2: The first of the incentive compatibility constraints that has to be met in equilibrium for player 1 is: if player 1 chooses to play C, but D gets played, the payoff from apologizing should be higher than the payoff from not apologizing.

$$w + \sum_{t=1}^{t=\infty} \delta^t E_{DD}^1 \leq (w - s) + \sum_{t=1}^{t=\infty} \delta^t (E_{CC}^1 - (1-p)s) \qquad (1)$$

The left hand side (LHS) of equation (1) is the payoff of player 1 from not apologizing once D is played out. The right hand side (RHS) of the equation is the payoff from apologizing in the same case. The timing of the game ensures that decision to apologize or not is made after the payoffs are realized and therefore the payoffs in t=0 are known with certainty. Further, the continuation payoff in the LHS is determined by the equilibrium strategy for player 2 which is to play D if there is no apology forthcoming after a (D,C) outcome. Player 1, knowing this, can't do any better than choosing to play D forever too. Equation (1) can be rewritten as (1.1), which leads to the first incentive compatibility constraint (IC).

$$(1-p)\delta s + (1-\delta)s \leq \delta[E_{CC}^1 - E_{DD}^1] \qquad (1.1)$$

IC 1: $s \leq \frac{\delta[E_{CC}^1 - E_{DD}^1]}{(1-p\delta)}$ for player 1 to be better off apologizing while choosing to play C.

This condition establishes the upper bound for the value of s. The cost of an apology should be therefore less than the discounted value of the difference between the cooperative outcome and the non-cooperative outcome. This result is quite intuitive and ensures that the cost of apology cannot be higher than the benefits of an apology.

Stage 1: In accordance with the 'one-stage deviation principle, the second incentive compatibility constraint that has to be met in equilibrium for player 1 is that player 1 must be better off choosing to play C and apologizing (as required by the proposed strategy) than choosing to play D and apologizing if D gets played.

$$\sum_{t=0}^{t=\infty} [E_{DC}^1 - ps] \leq \sum_{t=0}^{t=\infty} [E_{CC}^1 - (1-p)s] \qquad (2)$$

The LHS of equation (2) is the payoff of player 1 from choosing to play D and apologizing when D gets played, which happens with probability *p*. The RHS is the payoff from choosing to play C and apologizing when D get played, which happens with probability (1-*p*). Equation (2) can be written as (2.1), which leads to IC 2.

$$E_{DC}^{1} - E_{CC}^{1} \leq (2p-1)s \qquad (2.1)$$

IC 2: $s \geq \frac{[E_{DC}^{1} - E_{CC}^{1}]}{(2p-1)}$ for player 1 to be better off choosing to play C and apologizing than choosing to play D and apologizing.

This condition specifies the lower bound for the value of *s*. The cost of an apology should therefore be higher than the gains that player 1 might get from choosing to play D. IC 1 and 2 capture the trade-offs with respect to the cost of an apology. As these two conditions cannot be contradictory, it must be the case that in equilibrium the following condition should hold:

Lemma 1(Equilibrium Condition 1): $\frac{[E_{DC}^{1} - E_{CC}^{1}]}{(2p-1)} \leq \frac{[E_{CC}^{1} - E_{DD}^{1}]}{(1-p\delta)}$ for player 1 to play the proposed strategy in equilibrium.

4.2 Incentive Compatibility Constraint for Player 2

In addition to the condition posited in Lemma 1, the parameter for the rate of time preference δ, has to be large enough such that player 2 does not have an incentive to deviate[23] in stage 1. This implies:

$$E_{CD}^{2} + \sum_{t=1}^{t=\infty} \delta^{t} E_{DD}^{2} \leq \sum_{t=0}^{t=\infty} \delta^{t} E_{CC}^{2} \qquad (3)$$

This equation is based on the trigger strategy of player 1, which is, if player 2 plays D in any repetition, player 1 chooses to play D in all subsequent repetitions. Equation (3) can be written as (3.1), which leads to equilibrium condition 2:

$$E_{CD}^{2} + \frac{\delta E_{DD}^{2}}{1-\delta} \leq \frac{E_{CC}^{2}}{1-\delta} \qquad (3.1)$$

Lemma 2 (Equilibrium Condition 2): $\delta \geq \frac{E_{CD}^{2} - E_{CC}^{2}}{E_{CD}^{2} - E_{DD}^{2}}$ for player 1's trigger threat to be credible and for player 2 to play the equilibrium strategy.

[23] As player 1 cannot choose his actions with certainty: if δ large enough that player 2 does not deviate due to the threat of the grim trigger, it will be large enough to deter player 1 from choosing to defecting.

Thus, for the proposed sorry equilibrium to exist, both of these conditions have to hold[24].

4.3 Existence of Sorry Equilibrium

Proposition 1: For every *p,* there exists a δ that satisfies both EQ1 and EQ2. Therefore the proposed sorry equilibrium exists.[25]

The intuition for the existence of this equilibrium is simple: Both the players are better off when both of them choose to play C in every repetition. Now, in equilibrium, the cost of apology is so high that if player 1 does undertake the cost, it is sufficient to convince player 2 that defection is not intentional. This ensures that the grim outcome (both players choosing to play D), which serves to deter both players from intentionally defecting, is not triggered accidentally. The apology therefore, offers a resolution mechanism that paves the way for the players to continue choosing to cooperate despite accidental defections.

The apology described by this equilibrium is essentially a kind of truth claim and veracity of this truth claim depends on the cost[26] incurred to make it. This sorry equilibrium also belies the trend of using costly punishment only as a deterrence to coerce cooperation. So, while the threat of the off-equilibrium grim trigger outcome does play a role, it is insufficient due to the uncertainty inherent in the game. It is the apology, the in-equilibrium costly punishment that allows for resolution of the undesirable outcome (D,C). Moreover, it also captures all of the characteristics of an apology identified in section II. Apologies are used in equilibrium as their need is necessitated by residual stochasticity embedded in the stage game. This separation of intent and outcome is sufficiently general to affect a wide variety of activities and explains the frequency of the usage of apologies in diverse contexts and situations.

The multi-lateral nature of the model setting also exposes a not so obvious fact about the role of apologies as a social norm. While it is imperative for player 1 to apologize in case of an offence, if the cost of the apology is as prescribed by IC 1 and 2, it is also imperative for player 2 to accept it. Therefore, while sorry is a self-inflicted cost, communities that allow for an apology will always ensure that an adequate apology is accepted.

[24] The apology in the proposed equilibrium strategy is prescribed to convince player 2, that any defection is accidental. However, in-order for player 2 to continue the interaction, the payoffs from continued interaction must be larger than the loss suffered due to defection. But, as $E^2_{CC} > E^2_{DD}$ by assumption, this condition will always be met and therefore player 2 is always better off accepting the apology.

[25] Proof provided in the appendix

[26] Such costs can be incurred in a variety of ways: public declaration accepting the offence caused; spending time and effort to convince player 2; acts of self-sacrifice: a farmer might burn a part of crops, another might physically hurt himself etc..

The cost of apology in equilibrium is dependent on the parameters of the game and can therefore change to reflect the differences across circumstances. The IC 1 and 2 describe a range but they do not offer a point value for the cost of apology.

4.4 Cost of Apology

If the cost of sorry was a choice variable, whereby in stage 2 of the game, player 1 not only decides on whether to apologize or not, but also on the cost of the apology:

Proposition 2: In a sorry equilibrium, $s = \frac{1}{2p-1}[E_{DC}^{1} - E_{CC}^{1}] = (w - h)$

Proof: As an apology is a self-inflicted cost, player 1 would choose the lowest possible value of s such that it is just enough to convince player 2 that any infractions committed were not intentional. Therefore, the value of s in equilibrium must equal its lower bound derived in IC 2.

This result implies that player 1 needs to give up the entire extra payoff that he gets from accidentally defecting to convince player 2 that the defection was not intentional. Further, while the need for apology is necessitated by uncertainty, the cost of the apology does depend on the extent of the uncertainty. The cost of apology is therefore not mitigated by the fact that it was accidental. This, however, has a useful practical consequence. If the equilibrium does exist, the players do not need to be aware of the exact value of the p to play out the sorry equilibrium strategy.

The other interesting aspect of the cost of apology in equilibrium is that it does not depend on the harm caused to player 2 by player 1's mistake. This is contrary to the common prescription in the law and economics[27] literature (Hermalin, Avery, and Craswell 2007). In fact, in this setup the cost of the apology is smaller than expectation damages[28], commonly used in the law and economics literature. The consequences of this difference will be explored further in section V.

These interpretations of the model's results do not rely on nature of the payoffs involved. The payoffs need not be monetary and could be emotional too. The model does not distinguish between the two and implicitly assumes that they are interchangeable and valued equally. This has important consequences for how the cost of an apology is interpreted. In primitive communities an apology may not be measured in monetary terms but with regard to the shame and the loss of social honor or prestige that admitting an offence entails. Additionally, this assumption implies that an emotionally costly apology might work where the payoffs from the stage game are purely monetary and vice versa.

---

[27] The damage payments in law and economics literature are compensation mechanisms, where player 1 would pay player 2. An apology, however does not involve any transfers.

[28] Expectation damages= (h-l) and (h-l) > (w-h), as $2h$ > (w+l) by assumption

4.5 Comparative Statics of *p* and δ

EQ 1 and 2 are constraints that describe the relation between the two parameters *p and* δ in the sorry equilibrium. It would be reasonable to expect that an increase in *p*, or the certainty with which the action chosen by player 1 gets played would allow for lower δ sustaining the equilibrium. The EQ 2 constraint is in line with this expectation[29]. The lesser the chances of player 1 making a mistake, the more potent is the threat of the grim-trigger for player 2. However, EQ 1 mandates that a higher *p* require larger δ to support the equilibrium[30]. This result is counter-intuitive as lower chances of an accidental defection are expected to reduce the importance of the promise of future returns. This conundrum is an outcome of the constraint posed by IC 1. This constraint implies that if *p* is higher, ceteris paribus, player 1 has a greater incentive to choose to play C and not apologize. In equilibrium this increased incentive to not apologize is offset by greater concern for larger future returns accrued from apologizing.

5 Social Welfare Comparisons

5.1 Social Welfare in Sorry Equilibrium

The existence of the sorry equilibrium establishes that an apology can resolve the problems posed by residual stochasticity. However, if such a resolution is not socially beneficial, it might not be desirable. The net benefit to society can be evaluated by calculating the net social welfare in a sorry equilibrium, assuming a simple utilitarian welfare function.

Considering that the costs incurred to undertake an apology do not accrue to the offended parties, but are assumed to be lost of society, the social welfare is:

Net Present value of [Net payoff of player 1 in equilibrium + Net payoff of player 2 in equilibrium]=

$$\sum_{t=0}^{t=\infty} \delta^t [E_{CC}^1 + E_{CC}^2 - (1-p)s] \qquad (4)$$

[29] Proof in appendix

[30] Proof in appendix

Lemma 3: The Net Social Welfare in a sorry equilibrium (NSW-se)= $\frac{(1+p)h+(1-p)l}{(1-\delta)}$

The net social welfare depends partly on the harm caused to player 2 (l) during accidental defections by player 1. This is in contrast to the cost of an apology and emphasizes the importance of taking the social welfare effects of an equilibrium concept into consideration. Further, the net social welfare calculation is meaningful only when compared to a benchmark. A natural benchmark that offers itself is 0, implying a comparison to a situation where the players did not interact. The outcome from a sorry equilibrium must be at least as good as the no interaction condition to be a viable alternative. This requires:

$(1+p)h+ (1-p)l > 0$

It is apparent that p affects the extent of social welfare, which is increasing in p, ceteris paribus. This condition also illuminates that a sorry equilibrium is not socially desirable where l is very large compared to h. This makes for an interesting comparison with the individual rationality constraint of player 2:

$ph + (1-p)l > 0$

The higher weight on h in the social welfare constraint makes it obvious a sorry equilibrium can be socially desirable even if it is not individually beneficially. This result informs the basis for the implicit assumption in section IV that the players do not have the option to not play the game. However, in cases where an accidental defection might result in great harm like debilitating injury, death or large loss of property, it is not socially desirable to enforce a sorry equilibrium. However, not playing the game is not the only alternative, there are other community enforcement mechanisms and these may provide useful alternatives.

5.2 Social Welfare with Standard Grim-Trigger Strategy

The stand-alone grim-trigger strategy equilibrium is an often-used theoretical device and therefore offers a useful benchmark. The strategy prescribes that in case of any deviation from the (C,C) outcome, both players start playing (D,D). This works very well to enforce cooperation when both the players can choose their actions with certainty. In this setting, however, accidental deviations will happen and set off the grim trigger.

Given this strategy, the expected payoffs in each repetition of the stage game if both player 1 and 2 choose to play C, will be[31]:

$t=0: p(h+h)+(1-p)(w+l)$

$t=1: p^2(h+h)+p(1-p)(w+l)+(1-p)K$

$t=2: p^3(h+h)+p^2(1-p)(w+l)+(1-p^2)K$

[31] A rough proof that such a grim-trigger strategy can constitute an equilibrium is provided in the appendix

$t = 3: p^4(h+h) + p^3(1-p)(w+l) + (1-p^3)K$

and so on… (Where $K = E^1_{DD} + E^2_{DD}$)

The payoff at t=0 is just the combined expected payoffs of the two players. Due to the probability of accidental deviation, in every subsequent period t, the trigger is activated with the probability $(1-p^t)$, while the players continue with the cooperative strategy with probability p^t. However, in every period t an accidental deviation occurs with probability $p^t(1-p)$, increasing the chances of the trigger strategy being played in the next repetition.

Now, assuming the same utilitarian social welfare function, the social welfare from this strategy is:

$$\sum_{t=0}^{t=\infty} \delta^t p^t [p2h + (1-p)(w+l)] + \sum_{t=0}^{t=\infty} \delta^t (1-p^t) K \qquad (5)$$

Lemma 4: Net Social welfare from Grim-trigger (NSW-GT)=

$$\frac{[p2h + (1-p)(w+l)]}{1-\delta p} + K\delta\left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right]$$

Therefore, for the net social welfare in a sorry equilibrium to be higher than with standard grim trigger: (NSW-se) – (NSW-GT) > 0

This reduces to[32]:

$$\frac{(1-p)}{(1-\delta)1-\delta p}[h(1+\delta p) - w(1-\delta p) + \delta p2d] > 0 \qquad (6.1)$$

Proposition 4: The net social welfare in a sorry equilibrium higher than with standard grim-trigger strategies if: $h(1+\delta p) + \delta p2d > w(1-\delta p)$

The higher weight on h is on account of the continued interaction in a sorry equilibrium; the presence of w is due to the higher payoff for player 1 in case of accidental defection with no provision for an apology[33] and d captures the cost of the trigger strategy. Further, a higher p or δ, ceteris paribus, make a sorry equilibrium more attractive. A higher p has the dual effect of increasing the probability of the cooperative outcome being played out and reducing the expected cost of apology by risk and cost of accidental defections. A higher δ, implies more patient players, this also favors a sorry equilibrium as it offers continued cooperation in the long run.

This condition implies that if w is not much larger than h and/or d is sufficiently high, then a sorry equilibrium provides for higher social welfare. This implies that if the cost of an apology is low or alternately the cost of the trigger strategy (d) is high or both,

[32] Calculations in the appendix

[33] In a sorry equilibrium, the cost of apology ensures that the payoff for player 1 is h instead of w.

then a sorry equilibrium offers a better alternative. The effect of d on the difference in social welfare outcomes also offers insight into the paradox that using a standard grim trigger in a game with uncertainty leads to. A higher d makes the threat of the grim-trigger more credible. But, as the grim trigger outcome gets played in the equilibrium, a higher d reduces the social welfare from using such a strategy.

However, even though the standard-grim trigger might be useful in some cases, it does not provide a viable alternative to the sorry equilibrium if the cost of accidental defection (l) is very high.

5.3 Social Welfare under Ostracism

Ostracism is a practice that involves exclusion from social acceptance by general consent. It was practiced by the ancient Greeks and is still a part of the social norms in communities across the world (Williams 2002; Williams 1997; Posner 1997). The norm relies on the threat of exclusion from community to enforce acceptable behavior. If adapted to reflect the stage game modeled in section III, any player who defects would be ostracized from the community. It is also pertinent to point out that being ostracized from a community is likely to be very costly. It not only implies drastically reduced social interaction but also seriously hinders the ability of the ostracized individual to participate in economic activities.

The expected payoffs in each repetition of the stage game if player 2 plays C and player 1 chooses to play C, assuming that the long term costs of being ostracized is lumped together in one parameter O_s to be realized only in the period in which the player defected[34]:

$t = 0: p(h+h) + (1-p)(w+l-O_s)$

$t = 1: p^2(h+h) + p(1-p)(w+l-O_s)$

$t = 2: p^3(h+h) + p^2(1-p)(w+l-O_s)$

$t = 3: p^4(h+h) + p^3(1-p)(w+l-O_s)$

and so on..

The payoff at t=0 is just the combined expected payoffs of the two players minus the cost of player 1 being ostracized. In the second repetition, interaction continues with probability p and the cooperative outcome is played out only with the probability $p \times p$. Player 1 would defect with probability $p \times (1-p)$ in the second repetition. This pattern continues so that at time t the probability with which the cooperative outcome is played out is p^{t+1} and the probability for defection is $p^t(1-p)$.

Now, assuming the same utilitarian social welfare function, the social welfare from ostracism is:

[34] There is sketch proof in the appendix that such an outcome can be sustained as an equilibrium

$$\sum_{t=0}^{t=\infty} \delta^t p^t [p2h + (1-p)(w + l - O_s)] \quad (7)$$

Lemma 5: Net Social welfare from Ostracism (NSW-O)=

$$\frac{[p2h + (1-p)(w + l - O_s)]}{1 - \delta p}$$

Therefore, for the net social welfare in a sorry equilibrium to be higher than with Ostracism: (NSW-se) – (NSW-O) > 0

This condition reduces[35] to:

$$\frac{(1-p)}{(1-\delta)1-\delta p}[h(1+\delta p) + (1-\delta)(O_s - w) + l\delta(1-p)] > 0 \quad (7.1)$$

Proposition 5: The net social welfare in a sorry equilibrium higher than under ostracism if: $h(1+\delta p) + O_s(1-\delta) > w(1-\delta) - l\delta\,(1-p)$

This equation allows for a comparison of the trade offs involved in these two different types of social governance mechanisms. The costs involved in ostracizing players are captured by O_s. The costs involved in a sorry equilibrium are the cost of apology and the cost of accidental defection (due to continued interaction). So the social welfare from ostracism will the higher if w is much larger than h and/or l is very large and/or O_s is very small. Conversely welfare in a sorry equilibrium is to be higher if O_s is very large. Given this, barring a situation in which l constitutes a very large loss or the payoff from defection (w) being inordinately large; O_s is likely to be high enough to ensure that a sorry equilibrium leads to a better outcome. This is because O_s reflects the lifetime costs of ostracism, which include the social costs of isolation, the resulting humiliation and the foregone economic opportunities. Further, ostracism is not only costly to the individual; the social costs of ostracism also include the cost to the community from loosing a potentially economically productive member. For instance if the only washer-man in a community is ostracized, it could cost the community very dearly.

These costs incurred in the two systems also yields benefits. In the case of ostracism this benefit is that none of the players would deliberately defect. Further, if the damage from accidental defection is very high, ostracism ensures that such damages are not repeated in equilibrium. The cost this benefit is very high due a limitation that it shares with the standard grim-trigger: inability to distinguish between intentional and accidental defections.

The great benefit of a sorry equilibrium is that of continued interaction. Consider, the case of a hospital, even when the cost of accidents is very high (patients might die) the

[35] Calculations in the appendix

services offered are essential and the interaction must continue. The sorry equilibrium offers a mechanism to continue such interactions. However, the benefits from continued cooperation are not restricted to increased probability of getting a (C, C) outcome. Such dynamic benefits, while not captured in the stylized static model presented here, may have significant consequences. Continued interaction allows for learning, for development of processes that may in the context of this model, increase p or h, or reduce l over time. This is particularly relevant for the case of new activities (innovation), which involve a higher degree of uncertainty and if interactions that foster its improvement are not allowed for, they will likely never develop. For instance, if the first few iron forgers were ostracized for accidentally making poor quality iron, iron technology would never have developed. This process of potential improvement embodies the great benefits of having a governance mechanism that allows for apologies and therefore renewed engagement, even if one of the parties was short-changed.

5.4 Apologies: Joy of Receiving and Compensation

In the previous sections it has been shown that the proposed sorry equilibrium seemingly offers an advantage over more traditional methods of multi-lateral governance, under certain conditions. However, depending on the relative value of the payoffs, it comes up short in social welfare comparisons to mechanisms that do not account for the uncertainty inherent in the stage game. This is partly due to the fact that the model, stylized and austere as it is, does not account for all the benefits of unrestricted long-term engagement. But is it also partly due to the fact that the entire cost of apology is assumed to be lost to society. However, there exists anecdotal evidence to support the argument that this is not always the case. The legal literature on apologies, posits apologies as "contributing to the psychological health and well-being of the people involved" (Keeva 1999). There is also some evidence that "apology is a therapeutic balm" and "helps reduce the victim's anger" (Shuman 2000).

Considering that an apology offers some benefits over other forms of community governance, it can be argued that societies might work towards making it more viable. This could be achieved by inculcating preferences that allow for mitigation of the harm suffered by receiving an apology, a kind *of joy of receiving* preference. Such an idea is not entirely fanciful, the economic literature on pro-social preferences propounds: "Societies go to great lengths to instill such preferences in children during their process of socialization in families, school, and religious establishments, and continue the process in adults" (Dixit 2009). There is also evidence from the Japanese legal system that defendants offering an apology typically also seek to compensate[36] the victims they have harmed (Haley 1995; Stephens 2008).

These arguments imply that the entire cost of an apology may not always constitute a sunk cost to society. In some cases, a compensation amount maybe offered to the offended. In other cases an apology might indirectly increase the payoff of the player to whom the apology is being offered. Incorporating these claims into the sorry equilibrium, the cost of apology that is lost to society is: $(1-k)(w-h)$, where $k \in [0,1]$ is the proportion of the cost that is directly or indirectly transferred to the other player. While such a mechanism is not essential for the existence of a sorry equilibrium, it certainly improves its social welfare consequences, potentially making it more viable in a wider variety of situations.

5.5 Model Assumptions

In conducting this investigation of apologies the model proposed in section III makes certain strong assumptions. However, most of these are required to isolate the mechanics of the existence, usage and effectiveness of apologies from other related phenomena. For instance, residual stochasticity does not negate that some (probably a large) proportion of the stochasticity innate in activities is a function of the care taken in performing them. But a model that accounts for both must make an essentially arbitrary choice in deciding what proportion of such stochasticity is residual. More importantly, this choice could either over emphasize or obscure the role of apologies. Similarly, the model does not account for the dynamic benefits of continued interaction of the type mentioned in section IV. The extent of these benefits (which again would be an arbitrary choice) could easily bias the results of a social welfare analysis. Strong assumptions have also made about information that the players have about payoffs in all eventualities of the game. In everyday experience, however, this is not always true and points to the importance of beliefs. Different beliefs about the outcomes could lead to differing beliefs about the cost of apology potentially causing the equilibrium to break down. But beliefs are not accounted for in the model, as they are not fundamental to operationalize apologies. The additional insights that a more complex model might offer are traded for a more spartan exposition allowing for better tractability.

6 Courts

Some form of a formal legal system, involving courts, lawyers, judges and juries can be found in almost every country in the world. The laws that such systems are charged

[36] It is difficult to compare the use of apologies in courts directly to its use in community. However, this might one of the ways to ensure imposition of costly apology in courts.

with implementing vary considerably across jurisdictions and are often very complicated, riddled with many clauses, caveats and exceptions. This makes modeling a formal legal system almost impossible. Nonetheless, replicating a generic simplified form of court enforcement might yield useful insights; assisting in identifying limitations, complications and the possibility of perverse incentives. A useful method to proceed in such an investigation is to posit the ideal court and then proceed to add restrictions that reflect the real world and identify if the courts still fulfill their purpose. An alternative is to posit what a court is likely to do in a given case and then evaluate if it is optimal. Both these approaches shall be utilized in this section to evaluate the limitations of a formal legal system relative to a multi-lateral community governance mechanism with apologies.

Models of legal system in the economic literature (Cameron 1988; Nash 1991) mostly reflect and evaluate its efficacy in deterring infractions. This is so, even in the literature on tort law, where dealing with probabilistic events (Brown 1973), courts can deter accidents that might happen due to less than due care. The actions available to courts are: incarceration, compensation (transfer) and fines. Incarcerating offenders shares some of its characteristics with ostracism, so the investigation in this section will be limited to compensation and fines. An ideal court system would therefore, have sanctions that deter both players from deliberately defecting and when an infraction does occur, initiate some action to compensate and convince the offended player that the defection was accidental.

The easiest way to have such a court system would be to have an authority that can read intentions and inflict sanctions of its own accord[37]. The threat of sanctions would ensure that players do not intentionally defect and if needed, some kind of compensation mechanism can be instituted to placate the offended player. Such a system is of course fanciful but it does illustrate the need for more realistic assumptions about the legal system.

The first assumption has to be that courts (like players) cannot read the intentions of a player. But, this creates a conundrum as to what constitutes a crime, an act worthy of sanction/punishment: Should an accidental defection be punished, even when it is not possible to know if it is intentional or not? The role of intent in law is rather complicated and difficult to fully capture in a stylized setting. The simplifying assumption here shall be that while courts cannot distinguish between intentional and unintentional harm; an accidental harm is less likely to be punished. So if the probability of the court punishing

[37] Is it surprising then that kings, who also used to dispense justice, were frequently considered to be Gods in many parts of the world?

(finding the player guilty) an intentional defection is $\tau \epsilon$ [0,1], then the probability of an unintentional infraction being punished is τ', where $\tau' < \tau$. Further, any intervention by the courts requires that there be a plaintiff and defendant[38]. Both the defendant and the plaintiff are assumed to incur the same litigations cost c in the event of a suit being filed. The other assumptions are:

1. If the court rules in favor of the plaintiff, the defendant shall be required to pay damages D. These damage payments will be considered as transfer payments to the other player, unless mentioned otherwise.
2. The parameter $\tau \epsilon$ [0,1] reflects the probability with which the court will rule in favor of the plaintiff. It can also be thought of as the probability with which the plaintiff is able to satisfy the burden of proof (Masten, Scott E. and Prufer 2013). Further, the courts are restricted to making type-2 "False-negative" errors.[39]

The second assumption assures that a defecting player has no incentive to file a suit[40]. Further, a cooperating who files a suit against a defector has an excepted payoff of $(\tau D - c)$, while the defector will have an expected payoff $(-\tau D - c)$. These payoffs imply that for a cooperating player to have an incentive to file a suit, $\tau D > c$. However, compensating for costs of litigation are not the aim of damages, rather it strives to deter defection. Assuming the court deems that the minimum transfer amount to deter infractions is T, damages must be:

$$D = max\left\{T, \frac{c}{\tau}\right\}$$

6.1 Courts and Uncertainty

If a court does not account for probabilistic play it would consider the stage game as a standard prisoner's dilemma and decide the transfer amount in accordance with that assumption. Since, court enforcement cannot presume repeated interaction, so the transfer must be such that it is sufficient to dissuade defection in a one-shot game:

$$h \geq w - \tau T - c \qquad (8)$$

Where (8) can be written as (8.1):

$$T \geq \frac{w - h - c}{\tau} \qquad (8.1)$$

The transfer amount is set such that the payoff from cooperating be higher than from defecting. Consider the case where $\frac{c}{\tau} > T$[41]. In this case, even if the damage payments

[38] Having a single plaintiff and a defendant precludes the possibility of the players suing each other.
[39] This entire model of courts is a derivative of the model proposed in (Masten, Scott E. and Prufer 2013)
[40] As the courts do not make type 1 (false-positives) errors, if a defecting player files a suit he will merely incur a cost of $-c$.
[41] Where, $T = \frac{w-h-c}{\tau}$, the smallest transfer required to maintain cooperation. This result would obviously also hold for any larger T.

were sufficient to continue cooperation, the cost of litigation, $2c$ would be higher than the cost of an apology. Considering the social welfare impact, this would make a sorry equilibrium more viable[42].

Therefore, it is assumed that c is small enough such that $D = T$. This requires that courts be aware of the payoffs of stage game and the value of τ. Such an assumption might be unreasonable but it serves to show the limitations of the process, even with high availability of information.

The role of D is to deter defections and placate the offended player. It is clear that D is sufficient to deter player 1 from intentionally defecting[43] iff:

$$ph + (1-p)(w - \tau' D - c) \geq (1-p)h + p(w - \tau D - c) \qquad (9)$$

But, equation (9) requires that damages must be of the size given by (9.1) to deter player 1:

$$D^* \geq \frac{(2p-1)}{p(\tau' + \tau) - \tau'}[w - h - c] \qquad (9.1)$$

Now, if $\tau' = \tau$, $D^* = D$. But as this not the case, the damage payment decided on by the court might be too small to deter player 1 from deliberately defecting. This is because, comparing the RHS of equations (8.1) and (9.1):

$$\frac{(2p-1)}{p(\tau' + \tau) - \tau'} > \frac{1}{\tau} \qquad (10)$$

This is particularly troublesome as the difference between D^* and D would not be resolved if the court is aware of the residual stochasticity in action selection and is able to account for it in determining T. The nature of accidental defections, which makes them difficult to adjudicate on, would continue to adversely affect the efficacy of the court mechanism.

This model of courts and its resultant limitations might seem to be a caricature but it happens to reflect some of the elements of the Indian Judicial System. Historically, the Indian system has been given in to low compensation and fines[44] (Srinivasan and Eyre 2007; Galanter 1985). While there can be no denying that there are other institutional factors at play, insufficient damages, as this model shows, allows for intentional defection thereby increasing the probability of infractions. This insight lends to credence to the inference that given institutional capacity, higher number of infractions would severely limit the efficacy of courts and contribute to the log-jam that exists in Indian Courts (Hazra and Micevska 2004).

---

[42] Proof in appendix

[43] Even when intentionally defecting, the outcome will be (C,C) with probability p

[44] Fines for traffic offences from the Indian State of Bihar (1 EUR = Rs. 80 approx): http://transport.bih.nic.in/Penalties.htm

6.2 Role of intent in Jurisprudence and Large Damages

The most straightforward solution to the problem encountered in case I, is for courts to award large damages. The incentive compatibility constraint of player 1 is such that any $D > D^*$ would deter intentional defection. Once such compensation scheme often employed by courts are expectation damages (h – l). However, if the difference between τ and τ' is very small[45], even expectation damages may not be large enough. This is because the expected payoff without any damages from choosing to play D in greater than choosing to play C for player 1. Therefore, If the difference between τ and τ' is very small, it makes choosing to play D more profitable. In fact the smaller the difference, the closer a court mechanism is to a standard grim-trigger strategy or ostracism. Both of these social governance mechanisms fail to distinguish between an intentional and an accidental defection and if the difference between τ and τ' is small, the courts suffer a similar limitation. This observation relays another one of the reasons why intent should be a very important consideration in jurisprudence[46].

The other alternative before courts is to award very large damages[47]. But such big transfers entail large-scale redistribution and can potentially create perverse incentives. So even if these large damages can enforce continued cooperation, they might have undesirable side effects. Some of these side effects can be illustrated by relaxing some assumptions made in case I:

1. In case I, the courts were assumed to make only type 2 errors. However real courts are likely to make type 1 (false positives) errors too. If such errors are allowed, transfers larger than the total discounted value of the payoff from continued interaction (expectation damages would be larger than the discounted payoff for player 2 if l is large) would generate perverse incentives. Players might have incentives to sue if the outcome is (C,C) or even when they themselves have defected.
2. It courts can make type 1 errors, large transfers can seriously impact the community: individuals or their businesses can go bankrupt, hospitals shut down and companies become averse to experimenting with new products. Thus, higher compensation schemes might require courts to drastically reduce chances of error to be economically and socially sustainable. But this is likely to

[45] For a given τ, smaller τ'; $\frac{(2p-1)}{p(\tau'+\tau)-\tau'}$ decreases the denominator and increases the weight on (w-h-c)

[46] This also points to a potentially important role that confessions and associated reduction in damages should play in courts. However, a detailed discussion on the matter is avoided as the treatment of confessions in law is very context specific and differs significantly across jurisdictions.

[47] Another alternative is to fine the offender, but given the large fines required, this would lower the social welfare effects of courts.

increase costs of litigation, c. This would adversely affect the social welfare impact of court mechanism vis-à-vis the sorry equilibrium.

These results, while not explicitly derived, reflect some elements of the U.S judicial system. The ILR[48] reports: "America has the world's costliest legal system. As a percentage of our economy, U.S. legal liability costs are double those of the UK, three times higher than those in France and five times higher than those in Japan." These and other statistics provided in the introduction to this paper are representative of the kind of overuse and high litigation costs that these inferences imply.

6.3 Courts: Beyond Deterrence

This section has presented the difficulties that courts face in deterring intentional defection. However, in the model of interaction proposed in section III, deterrence was not the only feature required by a governance mechanism. In fact social governance mechanisms of ostracism and standard grim-trigger were shown to be incapable of cultivating sustained long-term interaction. In a sorry equilibrium, an apology resolves the problem by distinguishing between accidental and intentional defections. No further compensation or inducement is required for the offended player (player 2) to continue cooperation. So if courts can manage to deter intentional defections and decipher the intent behind the outcome, they too can foster long-cooperation. However, even if courts can overcome the problems highlighted in cases II and I; there is another potential limitation in the formal legal system.

Damage payments (sufficient to deter intentional defection) are enforced by the courts and received by the offended player with a probability of τ'. So with a probability of $(1 - \tau')$, the offended player receives no compensation or validation of his claim that an offence has been committed. What happens in instances when courts do not order payment of damages? If the joy of receiving preference is indeed fostered by communities, as postulated in section V, then it is likely that the interaction will cease to continue. In fact such circumstances might have a diametrically opposite effect on psychological health and anger than that the legal literature ascribes to apologies.

6.4 The U.S, India and Japan

There are obvious limitations to the evaluation of formal legal systems presented in this section. This is primarily because a complete theoretical evaluation of the courts is not within the scope of this paper. The objective instead is to make observations on the potential failings of such a system in the context of the proposed stage game. However, despite not being derived rigorously, these observations are lent credence by how

[48] http://www.instituteforlegalreform.com/issues/lawsuit-abuse-impact

closely they resemble the statistics from the two legal systems that have been referred to.

These observations also serve to juxtapose the potential outcomes generated in courts to those generated in a sorry equilibrium. Such a comparison in addition to exposing the limitations of courts also provides a justification for the use of apologies in the Japanese legal system. As these limitations also reflect, the problems faced by the legal systems in both the U.S and India; it may be prudent to heed calls in the legal literature for seriously looking into how the Japanese legal system integrates the use of apologies (Robbennolt 2003; White 2006; Petrucci 2002). A part of the function of any legal system is to act as a grievance redressal mechanism that enables continued, productive interaction. Residual stochasticity in choosing actions requires mechanisms that can distinguish between the intentional and the accidental. Costly apologies offer one a mechanism, which reduces the cost of litigation (apologies rely on self-identification), the need for large damage payments (avoiding undesirable side-effects) and most importantly allow for continuation of interactions.

7 Conclusion

"An apology is the superglue of life. It can repair just about anything."

Lynn Johnston

The U.S reels from having a legal system that costs the economy billion of dollars. In India, the courts have literally ground to halt. Compare these countries to Japan, where successful integration of apologies in the legal system that has led to fewer incarcerations and lower litigation propensity. In fact apologies are used frequently in everyday interactions the world over to mitigate and resolve conflict situations. They have even been put to use to in seemingly intractable situations like in the case of the Truth and Reconciliation Commission in post-apartheid South Africa. These and many other examples instigate the need for economic analysis of apology, to establish the extent to which apology is the ‘superglue of life’.

In order to evaluate the usefulness and efficacy of apologies, it might be illustrative to consider a counter-factual: A world in which everyone can choose their action with certainty[49]. In this world, the legal system and other social governance mechanisms like standard grim trigger and ostracism are very effective. In fact, they are so effective in deterring defection that this world does not need any ‘glue’ as nothing ever ‘breaks’. On the contrary, a world where players face uncertainty in choosing actions, undesirable

[49] This world does not have any other sources of friction in interaction like information asymmetries. This allows for the only one source of difference between this world and the one proposed in Section III.

outcomes cannot be avoided. Accidental defections caused by residual stochasticity require a mechanism to reconcile the players, to glue together what might be broken. This world, in so far as it more closely resembles the world we live in, requires an apology. Further, given the ubiquitous use of apologies, either our society exists in a constant state of dis-equilibrium or apology must be an equilibrium outcome. Section IV establishes that a sorry equilibrium can exist. For example, the relationship between a dry-cleaner and customer can go on uninterrupted (with a credible threat to defection) if the dry-cleaner can fully control how clothes are cleaned. If however, every once in a while the cleaning process (lapse of memory of problem with machine) leaves a blemish on the clothes, the relationship gets more complicated. However, If the customer realizes that it could be a mistake (given residual stochasticity), then an apology from the dry-cleaner can sustain the relationship. This apology must be costly though and might involve the dry-cleaner waiving off the fee or offering a free wash the next time. But the dry-cleaner would willingly bear these costs if the benefit from continued patronage by the customer is important enough.

The existence of such an equilibrium, however, does not preclude the possibility of interactions that may not be worth being 'glued'. For example: the relationship between a showman who shows off his skill with knives at fairs to entertain audiences and the person who has the knives thrown at him. A mistake in such a scenario could lead to the death of an individual. Given social welfare concerns, it is possible that in such cases, more indiscriminate social governance mechanisms might be more desirable. A social governance mechanism like Ostracism for instance would ensure that the game is never played again.

The possibility to account for the intent of the player also ensures that apologies do not suffer some of the limitations that courts might suffer from. A formal legal system cannot always distinguish between intentional and accidental defections. This creates the possibility of legal systems generating perverse incentives. In a legal system more like India's, such incentives are likely to encourage more intentional defections. In a legal system more like the one existing in the US, these are more likely to create an environment that encourages frivolous lawsuits. In general legal systems might have costly undesirable side effects. Given that apologies do offer a way to avoid these side-effects and that it is already widely used, it might be worthwhile considering how it can be better integrated into formal legal systems, a la Japan.

The efficacy of apology notwithstanding, in conducting this investigation the model proposed in section III makes certain assumptions. While these modeling choices are useful, they nonetheless reflect some of the limitations of this analysis. For instance, the effectiveness of an apology might reduce when used too frequently or that it is not very useful in the case one-off interactions. A model that accounts for residual and controlled (by the extent of care) stochasticity simultaneously might generate more insight into the margins at which apologies are useful. Or, a model that allows for the role of beliefs might predict when an apology is no longer sufficient to ensure continued cooperation. Alternately, a model that explicitly models the dynamic benefits of continued interaction might forecast the cases in which the social mechanism of apologies must be bolstered with legal intervention to benefit from the long-run benefits of sustained cooperation. These potential extensions to this model reveal directions for future research to develop a more complete model of apologies.

Appendix

Proof of Proposition 1:

Consider Lemma 1 (EQ1): $\frac{[E_{DC}^{1}-E_{CC}^{1}]}{(2p-1)} \leq \frac{\delta[E_{CC}^{1}-E_{DD}^{1}]}{(1-p\delta)}$

as $\frac{1}{2p-1}[E_{DC}^{1} - E_{CC}^{1}] = \frac{1}{2p-1}(pw + (1-p)h - ph - (1-p)w) = (w-h)$, EQ1 reduces to :

$(1 - p\delta)(w - h) \leq \delta[ph + (1-p)w - pd - (1-p)l]$

=> $\frac{(w-h)}{w-pd-(1-p)l} \leq \delta$ ....(1)

Now, as $\delta \in (0,1)$, the LHS must be less than 1 $\forall p$.

$(w - h) < w - pd - (1-p)l$

=> $-h < -pd - (1-p)l$

This inequality always holds as $h > 0$ and $d, l < 0$, by assumption. Therefore EQ1 allows for some $\delta \in (0,1)$.

Lemma 2 (EQ2): $\frac{E_{CD}^{2}-E_{CC}^{2}}{E_{CD}^{2}-E_{DD}^{2}} \leq \delta$

By assumption 5, $E_{CC}^{2} > E_{DD}^{2}$; therefore $\frac{E_{CD}^{2}-E_{CC}^{2}}{E_{CD}^{2}-E_{DD}^{2}} < 1$. EQ2 also allows for some $\delta \in (0,1)$.

Now as the LHS of both EQ 1 and 2 are strictly less than 1 $\forall p$. Therefore, for every given *p*, there must be some δ that satisfies both the conditions.

Effect of increase of p on EQ1 and 2:

Consider (1) a derivative of EQ1: a unit increase in p, decreases the denominator by (-*d*+*l*), as $0 > d > l$ by assumption. As the numerator does not change, but the denominator decreases, the LHS increases. This in turn requires higher δ to support the inequality.

Consider EQ2: a unit increase in p, changes the:

Numerator: $w - h - d + l$

Denominator: $2w - 2d$

Now, $2w - 2d > w - h - d + l$ as $w - d > l - h$. This implies that denominator increase faster than the numerator and therefore the LHS decreases and p increases, allowing smaller δ to support the inequality.

Standard Grim-trigger strategy equilibrium

This rough proof relies on restricting the choice of willing defection to t=0. So players must decide on whether to cooperate or defect only in the first repetition. Given this restriction, the equilibrium conditions are:

1) Payer 2 does not have an incentive to deviate. This gives rise to the same condition as EQ2:

$$E_{CD}^2 + \sum_{t=1}^{t=\infty} \delta^t E_{DD}^2 \leq \sum_{t=0}^{t=\infty} \delta^t E_{CC}^2$$

2) Player must be better of choosing to play C than D:

Payoff from choosing C: $\sum_{t=0}^{t=\infty} \delta^t p^t [ph + (1-p)\, w)] + \sum_{t=0}^{t=\infty} \delta^t (1-p^t) E_{DD}^1$

Payoff from choosing D: $\sum_{t=0}^{t=\infty} \delta^t (1-p)^t [(1-p)\, h + pw)] + \sum_{t=0}^{t=\infty} \delta^t (1-(1-p)^t) E_{DD}^1$

$$=> \frac{[ph + (1-p)w]}{1-\delta p} + E_{DD}^1 \delta \left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right] > \frac{[(1-p)h + pw]}{1-\delta(1-p)} + E_{DD}^1 \delta \left[\frac{p}{(1-\delta)(1-\delta(1-p)}\right]$$

$$=> \frac{(2p-1)(h - w(1-\delta))}{(1-\delta p)(1-\delta(1-p))} > E_{DD}^1 \delta \left[\frac{(1-\delta)(2p-1)}{(1-\delta)(1-\delta p)(1-\delta(1-p)}\right]$$

$$=> \delta \geq \frac{w-h}{w - E_{DD}^1}$$

As $h > E_{DD}^1$, both condition 1 and 2 allow some $\delta \in (0,1)$ to support the equilibrium

Alternate method for calculating Social welfare in Sorry Equilibrium:

$t = 0: p(h+h) + (1-p)(h+l)$

$t = 1: p^2(h+h) + p(1-p^2)(h+h) + (1-p)(h+l)$

$t = 2: p^3(h+h) + p(1-p^3)(h+h) + (1-p)(h+l)$

$t = 3: p^4(h+h) + p(1-p^4)(h+h) + (1-p)(h+l)$

and so on…$(w+l)$ – $(w\text{-}h)$ = $(h+l)$

$$=> \sum_{t=0}^{t=\infty} \delta^t p^t p2h + \sum_{t=0}^{t=\infty} \delta^t (1-p^t) p2h + \sum_{t=0}^{t=\infty} \delta^t (1-p)(h+l)$$

$$=> \frac{p2h}{1-\delta p} + \frac{(1-p)(h+l)}{1-\delta} + p2h\delta\left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right]$$

Calculating (NSW-se) – (NSW-GT)

$$\frac{p2h}{1-\delta p} + \frac{(1-p)(h+l)}{1-\delta} + p2h\delta\left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right] - \frac{[p2h + (1-p)(w+l)]}{1-\delta p} - K\delta\left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right]$$

$$=> \frac{(1-p)(h+l)}{1-\delta} - \frac{(1-p)(w+l)}{1-\delta p} + \left[\frac{\delta(1-p)}{(1-\delta)(1-\delta p)}\right][p2h - K]$$

$$=> \frac{(1-p)}{(1-\delta)(1-\delta p)}[(1-\delta p)(h+l) - (1-\delta)(w+l) + \delta(p2h - 2pd - (1-p)(w + l)]$$

$$=> \frac{(1-p)}{(1-\delta)1-\delta p}[h(1+\delta p) - w(1-\delta p) + \delta p2d]$$

Ostracism Equilibrium

This rough proof relies on restricting the choice of willing defection to t=0. So players must decide on whether to cooperate or defect only in the first repetition. Given this restriction, the equilibrium conditions are:

1) Payer 2 does not have an incentive to deviate. This leads to:

$$\frac{[ph + (1-p)l]}{1-\delta p} \geq p(w - O_s) + (1-p)(d - O_s)$$

$$=> p(h - (1-\delta p)w) + (1-p)(l - (1-\delta p)d) \geq -O_s$$

2) Player must be better of choosing to play C than D.

$$\frac{[ph + (1-p)(w - O_s]}{1 - \delta p} \geq \frac{[(1-p)h + p(w - O_s]}{1 - \delta(1-p)}$$

$$=> O_s(2p - 1)(1 - \delta) > (2p - 1)(w(1 - \delta) - h)$$

$$=> \delta \geq \frac{w - O_s - h}{w - O_s}$$

As O_s is expected to be very large, both conditions are satisfied by some $\delta \in (0,1)\ \forall p$.

Calculating (NSW-se) – (NSW-O)

$$\frac{p2h}{1-\delta p} + \frac{(1-p)(h+l)}{1-\delta} + p2h\delta\left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right] - \frac{p2h + (1-p)(w + l - O_s)}{1 - \delta p}$$

$$=> \frac{(1-p)(h+l)}{1-\delta} + p2h\delta\left[\frac{(1-p)}{(1-\delta)(1-\delta p)}\right] - \frac{(1-p)(w + l - O_s)}{1-\delta p}$$

$$=> \frac{(1-p)}{(1-\delta)(1-\delta p)}[(h+l)(1-\delta p) + p2h\delta - (1-\delta)(w + l - O_s)]$$

$$=> \frac{(1-p)}{(1-\delta)1-\delta p}[h(1+\delta p) + (1-\delta)(O_s - w) + l\delta(1-p)]$$

Comparing welfare impact of courts in case I to that of Sorry equilibrium:

Assuming that the court system can enforce continued interaction, the social welfare from such a system is:

NSW-CCI: $\sum_{t=0}^{t=\infty} \delta^t [E_{CC}^1 + E_{CC}^2 - (1-p)2c]$

NSW-se: $\sum_{t=0}^{t=\infty} \delta^t [E_{CC}^1 + E_{CC}^2 - (1-p)s]$

Where, $s = (w - h)$ and $c \geq \frac{w-h-c}{\tau}$.

Therefore, (NSW-se - NSW-CCI) > 0 because:

$(1+\tau)c \geq w - h$ and $\tau \epsilon$ [0,1]; implying $2c > (w - h)$.

Bibliography

Abreu, D. 1988. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56 (2): 383–96.

Ashkenazi-Golan, Galit. 2004. "Confession and Pardon in Repeated Games with Private Monitoring and Communication."

Bhaskar, V., and Ichiro Obara. 2002. "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring." *Journal of Economic Theory* 102 (1): 40–69. doi:10.1006/jeth.2001.2878. http://linkinghub.elsevier.com/retrieve/pii/S0022053101928789.

BJS. 2012. "TOTAL U.S. CORRECTIONAL POPULATION DECLINED IN 2012 FOR FOURTH YEAR."

Bowles, Samuel, and Herbert Gintis. 1998. "The Moral Economy of Communities: Structured Populations and Evolution of Pro-Social Norms." *Evolution of Human Behavior* 19: 3–25.

Brown, John Prather. 1973. "Toward an Economic Theory of Liability." *Journal of Legal Studies* 2: 323–49.

Bull, Clive. 1987. "The Existence of Self-Enforcing Implicit Contracts." *The Quarterly Journal of Economics* 102 (1 (Feb)): 147–60.

Cameron, Samuel. 1988. "The Economics of Crime Deterrence: A Survey of Theory and Evidence." *Kyklos* 41 (2): 301–23.

Cohen, A. D., Olshtain, E. 1985. *Comparing Apologies across Languages. Scientific and Humanistic Dimensions of Language*.

Compte, Olivier. 2014. "Communication in Repeated Games with Imperfect Private Monitoring Author (S):" *Econometrica* 66 (3): 597–626.

Darby, Bruce W., Barry R. Schlenker. 1982. "Children's Reactions to Apologies." *Psychology, Journal of Personality and Social* 43 (4): 742.

Denant-Boemont, Laurent, David Masclet, and Charles N. Noussair. 2007. "Punishment, Counterpunishment and Sanction Enforcement in a Social Dilemma Experiment." *Economic Theory* 33 (1): 145–67. doi:10.1007/s00199-007-0212-0. http://link.springer.com/10.1007/s00199-007-0212-0.

Dixit, Avinash. 2009. "Governance Institutions and Economic Activity." *American Economic Review* 99 (1): 5–24. doi:10.1257/aer.99.1.5. http://pubs.aeaweb.org/doi/abs/10.1257/aer.99.1.5.

Fischbacher, Urs, and Verena Utikal. 2013. "On the Acceptance of Apologies." *Games and Economic Behavior* 82 (November). Elsevier Inc. 592–608. doi:10.1016/j.geb.2013.09.003. http://linkinghub.elsevier.com/retrieve/pii/S0899825613001322.

Folkes, Valerie S. 1984. "Consumer Reactions to Product Failure: An Attributional Approach." *Journal of Consumer Research*, 398–409.

Fudenberg, D, and E Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica* 54 (3): 533–54.

Fudenberg, Drew, Levine David, and Eric Maskin. 1994. "The Folk Theorem with Imperfect Public Information." *Econometrica* 62 (5): 997–1039.

Galanter, Marc. 1985. "Legal Torpor: Why So Little Has Happened in India After the Bhopal Tragedy." *Tex. Int'l LJ* 20: 273.

Gibney, Mark (ed). 2008. *The Age of Apology: Facing up to the Past*. University of Pennsylvania Press.

Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade : The Maghribi Traders' Coalition." *American Economic Review* 83 (3): 525–48.

Haley, John Owen. 1982. "Sheathing the Sword of Justice in Japan: An Essay on Law without Sanctions." *Journal of Japanese Studies* 8 (2): 265–81.

———. 1995. *Authority without Power*. Oxford University Press.

———. 1998. "Apology and Pardon Learning From Japan." *Behavioral Scientist* 41 (6): 842–67.

Hazra, Arnab K., and Maja B. Micevska. 2004. "The Problem of Court Congestion: Evidence from Indian Lower Courts". 88. Discussion Papers on Development Policy.

Hearit, Keith Michael. 1994. "Apologies and Public Relations Crises at Chrysler, Toshiba, and Volvo." *Public Relations Review* 20 (2): 113–25. doi:http://dx.doi.org/10.1016/0363-8111(94)90053-1. http://www.sciencedirect.com/science/article/pii/0363811194900531.

Henrichson, Christian, and Delaney Ruth. 2012. "The Price of Prisons: What Incarceration Costs Taxpayers."

Hermalin, Benjamin E., W. Katz Avery, and Richard Craswell. 2007. "Contract Law." In *Handbook of Law and Economics, Vol. 1.*, edited by A.M. Polinsky and S. Shavell, 3–138. Amsterdam: Elsevier.

Ho, Benjamin. 2012. "Apologies as Signals : With Evidence from a Trust Game Apologies as Signals : With Evidence from a Trust Game." *Management Science* 58 (1): 141–58.

Keeva, Steven. 1999. "Does Law Mean Never Having to Say You're SORRY? Going to Trial over a Case Is Costly, Frustrating–and Can Perhaps Be Avoided with a Simple Apology." *ABA Journal*, 64–95.

Krishnan, Jayanth K., and C. Raj Kumar. 2010. "Delay in Process, Denial of Justice: The Jurispudence and Empirics of Speedy Trials in Comparative Perspective." *Georgetown Journal of International Law* 42: 747–84.

Lind, Jennifer. 2011. *Sorry States: Apologies in International Politics*. Cornell University Press.

Masten, Scott E., and Jens Prufer. 2013. "On the Evolution of Collective Enforcement Institutions: Communities and Courts". 11-013.

McQuillan, Lawrence J., Hovannes Abramyan, Anthony P. Archie, Jeffery A. Johnson, and Anna Erokhina. 2007. "Jackpot Justice: The True Cost of America's Tort System". San Francisco.

Nash, John. 1991. “To Make the Punishment Fit the Crime: The Theory and Statistical Estimation of a Multi-Period Optimal Deterrence Model.” *International Review of Law and Economics* 11 (1): 101–10.

Nytagodien, Ridwan, and Arthur Neal. 2004. “Collective Trauma, Apologies, and the Politics of Memory.” *Journal of Human Rights* 3 (4): 465–75.

Ohtsubo, Yohsuke, and Esuka Watanabe. 2009. “Do Sincere Apologies Need to Be Costly? Test of a Costly Signaling Model of Apology.” *Evolution and Human Behavior* 30 (2). Elsevier Inc. 114–23. doi:10.1016/j.evolhumbehav.2008.09.004. http://linkinghub.elsevier.com/retrieve/pii/S1090513808000925.

Ostrom, Elinor. 2000. “Collective Action and the Evolution of Social Norms.” *The Journal of Economic Perspectives* 14 (3): 137–58.

Ostrom, Elinor, James Walker, and Roy Gardner. 1992. “Covenants With and Without a Sword: Self Governance Is Possible.” *The American Political Science Review*, 404–17.

Petrucci, Carrie J. 2002. “Apology in the Criminal Justice Setting: Evidence for Including Apology as an Additional Component in the Legal System.” *Behavioral Sciences & the Law* 20 (4): 337–62. doi:10.1002/bsl.495. http://www.ncbi.nlm.nih.gov/pubmed/12210972.

Polinsky, A. Mitchell, and Steven Shavell. 1984. “The Optimal Use of Fines and Imprisonment.” *Journal of Public Economics* 24 (1): 89–99.

Posner, Richard A. 1997. “Social Norms and the Law: An Economic Approach.” *American Economic Review*, 365–69.

Reiter, Rosina Mrquez. 2000. *Linguistic Politeness in Britain and Uruguay: A Contrastive Study of Requests and Apologies*. John Benjamins.

Robbennolt, Jennifer K. 2003. “Apologies and Legal Settlement: An Empirical Examination.” *Michigan Law Review*, 460–516.

Rubinstein, A. 1979. “An Optimal Conviction Policy for Offences That May Have Been Committed by Accident.” *Applied Game Theory*, 406–13.

Scheff, Thomas J. 1998. “Community Conferences: Shame and Anger in Therapeutic Jurisprudence.” *Rev. Jur. UPR* 67: 97.

Schlenker, Barry R, and Bruce W Darby. 1981. “The Use of Apologies in Social Predicaments.” *Social Psychological Quarterly* 44 (3): 271–78.

Schlenker, Barry R. 1980. *Impression Management: The Self-Concept, Social Identity, and Interpersonal Relations*. Brooks/Cole Pub. Co. (Monterey, Calif.).

Shuman, Daniel W. 2000. “The Role of Apology in Tort Law.” *Judicature* 83 (4): 180–89.

Sobel, Joel. 2006. “For Better or Forever : Formal versus Informal Enforcement.” *Journal of Labor Economics* 24 (2): 271–97.

Srinivasan, Murugesan, and Mathew Jane Eyre. 2007. "Victims and the Criminal Justice System in India: Need for a Paradigm Shift in the Justice System." *Temida* 10 (2): 51–62.

Stephens, MA. 2008. "I'm Sorry: Exploring the Reasons Behind the Differing Roles of Apology in American and Japanese Civil Cases." *Widener L. Rev.* 21 (2005): 185–204. http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/wlsj14&section=8.

Times of India. 2013. "Over 3 Crore Cases Pending in Courts in India: Govt." *The Times of India*.

TOI. 2010. "Courts Will Take 320 Years to Clear Backlog Cases: Justice Rao." *Times of India*.

Trouillot, Michel-Rolph. 2000. "Abortive Rituals: Historical Apologies in the Global Era." *Interventions* 2 (2): 171–86.

Viscusi, Kip W., and Michael J. Moore. 1993. "Product Liability, Research and Development, and Innovation." *Journal of Political Economy* 101 (1): 161–84.

White, Brent T (The University of Arizona). 2006. "Say You're Sorry: Court-Ordered Apologies as a Civil Rights Remedy."

Williams, Kipling D. 2002. *Ostracism: The Power of Silence*. Guilford Press.

Williams, Kipling D. 1997. "Social Ostracism." In *Aversive Interpersonal Behaviors*, 133–70. Springer US.