

Why do apology laws fail

Obdulio Varela

Abstract

Should apologies be admissible into evidence as proof of fault in civil cases? Between 1999 and 2011, the number of US states with *apology laws*, legislation that excludes the admissibility of apologies in court, increased from 2 to 38, along with all the Australian jurisdictions, the United Kingdom, most of the Canadian provinces, and Hong Kong. Legislators hope is that by passing these laws civil disputes will be settled more often, and lawsuits will be prevented. However, recent evidence from US shows that these laws have had the opposite effect: apology laws have increased the number of lawsuits. In this paper we provide an explanation for why apology laws fail. We model the decision of defendants to apologize and plaintiff to accept it at the shadows of the courtrooms when apologies are admissible and inadmissible at court. We show that apology laws may reduce settlement by encouraging insincere apologies that in turn makes plaintiffs to not accept it. We contribute to show on which type of relationships apology laws fails: apology laws precludes the settlement of cases that are socially valuable to be settled. We uncover two channel through which apology laws induce insincere defendants to apologize more often: i) by cheapening apologies, as now apologies cannot be used as evidence; and ii) by increasing the average damage awards, as now the judge cannot observe at trial rejected sincere apologies, she pools these cases with others. Moreover, when the unintended effect arises, apology laws induce more miscarriages of justice.

JEL Codes: K13, K32, K41, I18

1 Introduction

Should apologies be admissible into evidence as proof of fault in civil cases? In other words, in a civil case where Smith had sued Jones, and Jones has apologized for what happened, should we allow Smith to use Jones's apologies as evidence of Jones's fault at trial? This question has been debated in the legal arena for more than 20 years. The answer seems negative in the light of what happened in the following decades in common law jurisdictions: Legislation that limit the admissibility of apologies in court for proving civil liability, broadly referred as *Apology laws*, has become the norm. Between 1999 and 2011, the number of US states with apology laws increased from 2 to 38, along with all

the Australian jurisdictions, the United Kingdom, most of the Canadian provinces, and Hong Kong (Vines [2015], Myers [2016], Kleefeld [2017]). But why they have become so famous? As the administration of civil justice is increasingly compromised by high litigation cost and delays, legislators hope is that by passing these laws civil disputes will be resolved more amicably and less expensively or, in other words, they expected to prevent lawsuits and speed up settlements¹. Research in psychology, law and social sciences shows that an apology is a strong and cheap device to restore social or economic relationships that have been disturbed. Drawing from these results, legislators passed these laws with the following intended mechanism in mind: First, by “protecting” apologies, defendants will apologize more often; and second, given this proliferation of apologies, plaintiff will be ready to accept it.

Despite legislators’ intentions, recent empirical evidence shows that apology laws has the opposite effect: they decrease out-of-court settlements and increase cases at trial (Ho and Liu [2011], McMichael et al. [2019], McMichael [2021]). In this paper we offer an explanation for why and when apology laws fail. We use a signaling model to analyse the decision of defendants to apologize and plaintiffs to accept it at the shadows of the courtrooms on both institutional arrangements, and use the model to examine the effect of making apologies inadmissible on settlement/trial outcomes. We show that apology laws have the *intended effect* -more settlements/less trial- but also an *unintended effect* -less settlements/more trial. We show that apology laws may reduce settlement by encouraging insincere apologies that in turn makes plaintiffs to not accept it. Importantly, which effect dominates depends on the type of relationship considered and the context of the conflict. We show that the unintended effect of apology laws arises in relationships where parties benefit more if the relationship is restored -long term relationships-, and in context where the defendant’s reputation concerns are not too high. On the contrary, the intended effect of apology laws arises for comparatively less valuable -short term- relationships, and higher reputational concerns.

Our argument is based on a two-sided incomplete information environment. On the one side of the conflict, the plaintiff has incomplete information about the sincerity of an

¹“One new strategy to promote the early, effective and affordable resolution of disputes that is being considered by the Ministry of Attorney General is apology legislation.” (Discussion paper on apology legislation, British Columbia Ministry of Attorney General (2006), p.1). These laws seek to “reduce lawsuits and encourage settlements” (California Assembly Commission on Judiciary, TENN. R. EVID. 409.1 cmt.).

apology. In this sense, drawing from previous research on apologies, we allow for sincere and insincere apologies². We consider *sincere apologies* as a communication device designed to restore a broken relationship. In our framework, sincere and insincere defendants are differentiated by their respective gains from restoring the relationship. The latter ones issue *insincere apologies*, tendered just to avoid trial³. On the other side of the conflict, the defendant is not sure if an apology would be accepted. A subset of plaintiff will be ready to accept sincere apologies, i.e. the ones that comes from sincere defendants; but they will not accept apologies tendered just to avoid trial, the ones issued by insincere guilty defendants. Moreover, apologizing is risky: not all plaintiff are ready to accept sincere apologies. We allow for plaintiffs that do not care about re-establishing the relationship, and only want indemnity payment. They may use apologies at trial if admissible⁴. In this context, we show that apology laws reduce settlement by encouraging insincere apologies that in turn makes plaintiffs to not accept it. We uncover two channel through which apology laws induce insincere defendants to apologize more often: i) by cheapening apologies, as now apologies cannot be used as evidence in courts; and ii) by increasing the average damage awards, as now the judge cannot observe a rejected sincere apology at trial, she pools these guilty defendants with innocent ones.

We show that the unintended effect of apology laws arises in context where parties could benefit more if the relationship is restored. Before apology law, apologies were costly because they can be used in court if rejected. This potential cost of apologies couple with innocent defendants that refused to apologize, enabled plaintiff to infer the sincerity of an apology. On the one hand, sincere defendants with high enough valuation of the relationship took the risk of apologizing. On the other, insincere defendants preferred to remain silent, as they were pooled with innocent defendants and paid less. We show that

²There is a large body of evidence showing that a harmdoer who sends an apology is much more likely to be forgiven than a non-apologizer ([Weiner et al., 1991],[Ohbuchi and Sato, 1994],[Ohtsubo and Watanabe, 2009], [Abeler et al., 2010], [Fischbacher and Utikal, 2013], [Ho and Liu, 2011]). However, lab experiments showed that apologies can sometimes backfire if they are seen as insincere (Fischbacher and Utikal 2013, Gilbert et al. 2017)

³Cohen [2002], p.872: “The main arguments against these laws are that such laws could: (1) induce insincere, manipulative apologies from unremorseful injurers”.

⁴The model captures the vicious cycle that Benjamin Ho explains at *freakonomics*: “...where doctors are afraid to apologize because they’re scared of getting sued. But the patients, the only reason they sue is perhaps because they never got an apology. To combat this, a lot of states started passing what are called ‘I’m sorry’ laws” (<https://freakonomics.com/podcast/apologies/>, accessed on 25/09/2020).

apology laws reduce the settlement of cases that are socially valuable to be settled. On the contrary, apology laws may induce more settlement for cases where future gains derived from the relationship are not so high. The potential cost of using apologies in court is what prevented settlement in these cases, it was too costly even for sincere defendants, and the exemption of apologies may work.

Our results offer an explanation for the existing evidence of the unintended effect of apology laws. When a lawsuit is litigated fully, there is a public record of the events. In contrast, apologies often lead to “non-lawsuits” and private settlements about which data is more difficult to obtain. McMichael et al. [2019] shows the unintended effect of apology laws in the context of medical malpractice cases through a dataset that includes information on claims that resulted in lawsuits as well as those that did not –the ones that settled⁵. The data comes from a national malpractice insurer that contains information on 90% of all US physicians practising in a single specialty from 2004 through 2014. For the particular specialty they have data, some physicians focus primarily on seeing patients in an office setting, non-surgeons, while others both see patients in an office and perform surgery, surgeons. They found that apology laws do not have effect on settlement/trial outcomes for surgeons, but reduces settlement –and increase lawsuits– against non-surgeons. This is consistent with an explanation based on the physicians’ valuation of the relationship. In their database non-surgeons treat patient over the course of year, whereas surgeons interact in discrete events⁶. In this sense, our theory predicts that the unintended effect of apology laws arises for cases where physicians have high gains from restoring the relationship, provided apologies are not so costly. This explains the result for non-surgeons. On the contrary, when this valuation is low, as for surgeons, these laws have no effect.

Our predictions also explain the results of a recent study that uses novel information

⁵McMichael et al. [2019], p.341: “This dataset includes substantially more information than is publicly available and, thus, presents a unique opportunity to understand the effect of apology laws on the entire litigation landscape in ways that are not possible using publicly available data.”. Ho and Liu [2011] finds mixed evidence of the effect of apology laws in the medical malpractice context using publicly available data. This data only includes cases that ended with a positive payment to the claimant.

⁶McMichael et al. [2019], p.369:“Surgeons generally interact with and treat patients in a discrete event, i.e., the surgery they are performing plus any pre-operative and post-operative care. Because of this discrete interaction, patients who suffer an injury will likely have little trouble tracing that injury to an error that occurred during surgery. On the other hand, non-surgeons generally treat their patients over the course of years or may interact with patients a number of times when attempting to resolve an injury or illness. Thus, observing the malpractice of non-surgeon physicians may be more difficult.”

about medical malpractice insurance premiums charged to physicians across three separate specialties: general surgery, internal medicine, and obstetrics/gynecology⁷. McMichael [2021] shows that apology laws has caused an increase in malpractice premiums for all specialties, but the effect is higher for internists. As the focus of this specialty is on dealing with long doctor-patient relationships, our results provides an explanation⁸.

These findings have implications for legislators. Apology laws' intended effects are dependant on the type of broken relationship. In particular, the doctor-patient relationship may enter in the type where apology laws reduce possible settlement. This is relevant as in US 28 over 38 apology laws were approved for healthcare providers. We also show other negative outcomes of apology laws. First, we show that if there is a small cost to apologize for defendants when they know that an apology will be rejected, apology laws also reduce communication. This is important as apart from the pretended effect on litigation, legislators were also motivated by the idea that apology laws would encourage more apologies, that has intrinsic social benefits per se⁹. Second, we show that when the unintended effect arises, apology laws incur in more miscarriages of justice than before. Finally, our model uncovers two other negative factors about apology laws that were already raised by the literature. First, when apology laws have the intended effect, public confidence in the courts could be adversely affected. Second, if frivolous claims are prevalent, apology law encourage the acceptance of insincere apologies¹⁰.

⁷McMichael [2021], p.4 footnote 22: "...the data analyzed here offer two important advantages: (1) malpractice premiums represent an amalgamation of the factors that influence malpractice liability risk compiled by insurance companies, whose profitability depends on accurately capturing this risk, and (2) these data are not subject to the substantial problem of missing information that affects publicly available malpractice claims data."

⁸See AMERICAN COLLEGE OF PHYSICIANS, About Internal Medicine, <https://www.acponline.org/about-acp/about-internal-medicine>, detailing the care provided by internists: "Their training uniquely qualifies them to practice primary care and follow patients over the duration of their adult lives and establish long and rewarding personal relationships with their patients.", "...a discipline focused ... in caring for patients in the context of thoughtful, meaningful doctor-patient relationships." (accessed on 25/09/ 2020)

⁹"Factors in favour of apology legislation include: a.To avoid litigation and encourage the early and cost-effective resolution of disputes; b. To encourage natural, open and direct dialogue between people after injuries; and c.To encourage people to engage in the moral and humane act of apologizing after they have injured another and to take responsibility for their actions"(British Columbia Ministry of Attorney General (2006), , p.4).

¹⁰"Negative factors include: a. Public confidence in the courts could be adversely affected if a person who has admitted liability in an apology is found not liable; b. Insincere and strategic apologies could be encouraged; and c. Apologies encouraged by such legislation might create an emotional vulnerability

Our work is related to an existing apology literature in psychology, sociology, medicine, and law. In economics [Ho, 2012] introduces a costly signaling model to study apologies in a principal-agent framework where apologies are exogenously costly. Apart from that fact that our model has two audiences, a judge that make apologies endogenously costly, there is a difference in terms of the interpretation of apologies, admissions or excuses, that we discuss in section 4. Our work is also related to the literature that studies the outcomes of institutional rules that precludes particular evidence to be used at trial. [Daughety and Reinganum, 1995] studies the rule that exempt settlement offer on settlement outcomes, [Seidmann, 2005] the effect of the right to silence on convictions rate and accuracy. As in [Daughety and Reinganum, 1995] we analyse signaling games with two audiences, and compare the outcomes of making the signal private or public. In this sense, our model contributes to the literature of strategic information transmission with multiple audiences. Finally, there is a recent literature that focuses on how firms should apologies (see Halperin et al. [2019] for a large-scale field experiment with Uber’s customers and Abeler et al. [2010] for a field experiment with ebay). Our model can be cast in terms of a history of customer dissatisfaction with the service provided by a firm’s employee or seller. We can consider that the firm, or online platform, following a customer dissatisfaction with the employee/service, wants to encourage direct agreements or reconciliation (settlement) between a customer/buyer (plaintiff) and an employee/seller (defendant). Moreover, we can expect that the higher the gains from restoring the relationship -longer term relationship- the better for the firm’s perspective. In this context, the admissibility of apologies in courts can be translated to the verifiability of apologies by the firm. Our results highlight that the design of a communication system between customer/employee where apologies can be verified by the firm if tendered, as in ‘this communication may be recorded’ would improve firm’s objectives¹¹.

The work is organized as follows. In the next section we provide an overview of the

in some plaintiffs who may accept settlements that are inappropriately low” (British Columbia Ministry of Attorney General (2006), p.4). Cohen [2002], p.856 : “Sophisticated defendants are going take advantage of naive injured parties through these laws. They’ll issue apologies knowing that there’s no real risk involved, but naive injured parties will think these apologies are meaningful - that they do involve risk. Injured parties will think the injurers are putting their necks on the line when in fact they aren’t.”

¹¹In 2018 Uber introduced UberChat where customers can communicate directly with drivers (<https://eng.uber.com/one-click-chat/>, accessed, September 2020). Our results indicate that Uber should encourage drivers to apologize in cases they fail with frequent customers.

main elements of apology laws. In section 2 we present the model. In section 3 we present the analysis of the effects of apology laws. In section 4 we discuss our results with previous work in more detail.

1.1 Apology laws: how they work, and for which cases

Apology legislation includes statutory provisions that remove legal disincentives to offering an apology in the context of civil disputes. The regulatory framework differs to a large extent between various jurisdictions. The points of difference concern the type of apology that is protected, the interference with other areas of law (evidence law as well as substantive law) and the scope of coverage. We explain each of these features below. Despite the disparity in the landscape of apology legislation, there has been a trend towards a broader legal protection and a wider scope of application. The first apology act enacted in Massachusetts in 1986 provided solely for inadmissibility of benevolent statements as evidence of fault or liability in the context of an accident. While US states mainly followed this path (with some alterations), the introduction of apologies acts in the Australian states and territories (as of 2002) and in the Canadian provinces (as of 2006) signified a remarkable expansion of the scope. A large part of those acts apply to civil liability of any kind, cover apologies which may include admissions of fault and sometimes even encompass specific provisions on insurance and limitation. The Hong Kong Apology Ordinance of 2017 is described as “the most ambitious apology law yet”¹².

Legal effect: How they work. Excluding apologies from admissibility into evidence is one of the most widespread forms of evidentiary adjustments. Almost all jurisdictions limit the admissibility of apologies for proving fault or liability¹³. Inadmissibility means that apologies cannot be part of the decision-making process. A court will not receive nor consider them in deciding over liability. When there is a jury trial, the apologies shall not be available to jurors. The laws in most jurisdictions stipulate that apologies are not admissible “as evidence of fault or liability in connection with the matter” or something similar. From a common law perspective, an admission is an oral or written statement or conduct

¹²Nevertheless, there is no unequivocal pattern, as the aspirations of the Scottish Apologies Act of 2016 and the amendment of the Irish Civil Liability and Courts Act of 2015 are much more moderate. (see Kleefeld [2017] for a survey).

¹³The exception are England and Wales

made directly by or on behalf of a party which goes against the interests of that party. Admissions may be formal or informal. Apologies made outside the court room are associated with informal admissions. Unlike formal admissions, they are tendered as substantive evidence by the opponent, subject to contradiction or explanation and may be denied, explained away or contradicted by the maker (e.g. by establishing that it was made for some secondary reason and not true). Legislation in almost all jurisdictions stipulates in more or less similar terms that an apology does not constitute “an express or implied admission of fault or liability”.

Some jurisdictions have made the policy choice to broaden the act’s ambit beyond the amendment of evidence laws. The apology acts in the Canadian provinces, Hong Kong and Ireland expressly bars using an apology to void or otherwise affect insurance coverage¹⁴.

Definition of apology: full vs partial apology laws. The ‘law and apology’ literature often makes a distinction between full and partial apologies. Accordingly, the basic distinction mostly comes down to the question whether the legal provisions merely protect expressions of sympathy and benevolence, defined in the literature as “partial apology laws”, or also include fault-admitting apologies, “full apology laws”¹⁵. The former category provides for protection of expressions of a general sense of sympathy or benevolence. It encompasses statements such as “I am sorry this happened to you”. The latter refers to apologies which are not confined to expressions of sympathy or regret, but also incorporate admissions of fault or wrongdoing. From a legal perspective, such an apology is much more burdensome than a mere communication of sympathy. Australian and Canadian legislations underline this point by adding to the definition “whether or not the apology admits or implies an admission of fault in connection with the matter”. In six US

¹⁴Insurers would regularly tell their clients that an apology will render insurance coverage void. This would sometimes even be reflected in insurances clauses hindering insured persons from making apologies to those whom they injured.

¹⁵However, in accordance with apology theorists, an acknowledgement of fault without any expression of sympathy or regret should also be considered a partial apology. According to these scholars, apologies consist of different building blocks: an affirmation or acknowledgement of fault; an expression of regret, remorse or sorrow; a willingness to repair and a promise to adapt future behaviour. In this context, reference is also made to the four “Rs” (responsibility, remorse, reparation and resolution). Whereas partial apologies would consist of some, but not all of these components, full apologies would enclose all or at least the majority of them. Despite all jurisdictions have defined the term apology in their legislation, the willingness to repair (the action component) and the commitment to change future behaviour (the articulation of forbearance) are almost always lacking.

states, the definition of apology extends to admissions of fault as well¹⁶.

Scope of coverage. There is a diversity in scope of application of apology legislation as well. The Massachusetts act in 1986 provided solely for inadmissibility of benevolent statements in the context of accidents. While some US states have followed this path, most of them have even narrowed down the scope to medical malpractice cases. Other common law systems refer to civil liability of any kind¹⁷.

2 The Model

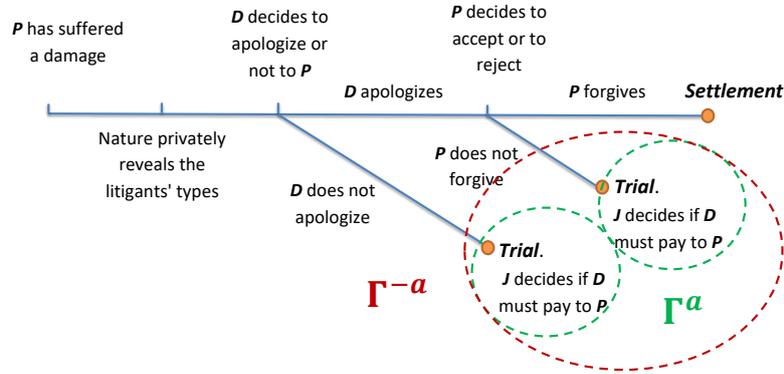
The objective is to characterize equilibria of two games: Γ^a , in which apologies are admissible in court, and Γ^{-a} , in which they are inadmissible. There are three players in both games: a defendant (D), a plaintiff (P) and a judge (J). The game begins after P has suffered damages. We consider the following timing for the analysis. After Nature has privately revealed the litigants' types, D decides whether or not to apologize to P , an action which J does not directly observe. If D does not apologize then the case goes to trial; otherwise P decides whether to accept or to reject it. If she accepts it, then the litigants settle; if she rejects it, the case proceeds to trial¹⁸. At trial, J decides whether to award damages of $m \in \{0, M\}$ to P : where $m = M$ ($m = 0$) implies that D was found (not) liable for the damages. J observes whether an apology has been made in Γ^a , but not in Γ^{-a} , if presented in court. This difference fully distinguishes the two games. Moreover, there is an exogenous probability, $0 < \lambda < 1$, that the trial itself reveals the true liability of D in the

¹⁶This happens in a more subtle way than under Australian or Canadian law. Definitions of apologies in the US normally consist of a list of emotions which are expressed or affirmed (i.e. apology, sympathy, commiseration, condolence, compassion, regret or a general sense of benevolence). Those six US states integrate terms such as fault, mistake and error or responsibility and liability into that list.

¹⁷This path is taken by Australian and Canadian lawmakers. Canadian includes generally intentional torts. A limited number of US states also refers to "any claim growing out of the event" or "a cause of action in tort". The UK Compensation Act applies to "negligence or breach of statutory duty". The Scottish Apologies Act and Hong Kong Ordinance even have a broader scope, covering "all civil proceedings" or "applicable proceedings". These apology acts can potentially impact other areas of law (such as family law, contracts, commercial litigation and administrative law).

¹⁸The reader can interpret the trial outcome as a trial/settlement bargaining outcome, as the settlement bargaining will be determined by the judge equilibrium decision. This simple extensive form facilitates the exposition of our results. In the appendix we show that they are robust to a settlement bargaining extension. Note that our interest is on settlement induce by apologies, not by bargaining arguments.

case, whereas with probability $1 - \lambda$, J must infer the liability by the equilibrium strategies of players. The next figure shows the timeline of both games.



The two litigants are each privately informed. D has 3 possible types, denoted by d : d_i (the innocent type), d_s (the sincere guilty type) and d_n (the insincere guilty type). The prior probability of type d_x is denoted by p_x . Now we describe the payoff for each d . First, every type loses m if she pays damages of m at trial. Second, we consider the payoffs associated with the action of apologizing and its consequences. We treat apologies as confession, that is, an apology communicates that a fault has been made. In this sense, we assume that an innocent type never apologizes, she will not provide a false confession. We discuss this assumption in the final section. For guilty types we allow for apologies to be costly –besides the potential legal consequences of apologies regarding liability– and this cost depends on whether apologies are accepted or not. When an apology is accepted, and parties settle, we denote the cost by t . We interpret t as a monetary transfer to compensate P and restore the relationship. On the other hand, when an apology is rejected, we denote the cost of apologizing by $c \geq 0$. This can be interpreted as a “reputation” cost for the defendant, a psychological cost of admitting fault, or the risk of future litigation with others potential damaged parties involved in the incident¹⁹. Finally, we allow for future payoffs

¹⁹Evidence shows that transgressors who apologize in situations in which competence is relevant suffer a negative impact on their perceived competence ([Kim et al., 2006], [Weiner et al., 1991]), and speakers are aware of this. Physicians are reluctant to give apologies because it damages their reputation (Gallagher et al. [2003]). See Chaudhry and Loewenstein [2019] for a theory, and evidence, that explains why apologizing

derived from a restored relationship. In line with research on apologies, an apology may serve to restore broken and possibly fruitful relationships. We denote this payoffs by $v_d \in \{v_s, v_n\}$, and it is what differentiates the sincere and insincere types. While d_s receives a possibly positive payoff, $v_s \geq 0$, if she apologizes and is forgiven; d_n does not care about restoring the relationship, $v_n = 0$, and so the only reason for her to apologize is to avoid trial. The ex post payoffs of the D 's types in the three possible outcomes are summarize below:

$$d\text{'s payoffs: } \begin{cases} v_d - t & \text{if she apologizes and is forgiven (Settlement)} \\ -c - m & \text{if she apologizes and is not forgiven (Trial)} \\ -m & \text{if she not apologizes (Trial)} \end{cases}$$

On the other hand, P has two possible types: P_F (the forgiving type) and P_L (the litigious type), with the prior probability of type P_L denoted by q_L . Neither type will litigate provided she knows that D is innocent. We introduce the possibility of frivolous claims later. Now we describe both types payoffs:

1. P_L does not care about apologies, she is interested in indemnity payments. Her payoff equals the damages she is awarded, m , irrespective of whether D has apologized, and will always take the case to trial whenever there is some probability that D is liable.
2. P_F is willing to forgive just sincere apologies, defined as the one that comes from d_s (not insincere ones, the ones tendered by d_n). Formally, in case P_F accepts an apology from d_s –and hence settles– she earns $v_F + t$: $v_F \geq 0$ represents the value of a restored relationship with d_s , and t the compensation received. In case P_F accepts an apology from d_n , she earns just t . Whenever she rejects an apology, or she ends at trial without having receive one, she earns the award m .

Finally, J wants to award her best guess about the damages given all admissible evidence. We assume that she gains zero when a correct award is sentenced, whereas the loses

is not costless. We consider that a trial process risks D to a higher level of public exposure in comparison to a settlement agreement, that is why we choose to not include c when an apology is accepted. Our results holds if we introduce differential cost by type of defendants (c_s and c_n). The reader can also interpret that this reputation cost is included in a negative trial sentence $c \subseteq M$.

incurred by court mistakes are symmetric: she loses the same amount either when a guilty D is exempted from indemnity payments or when innocent ones pay it. J 's loss if the case does not go to trial is irrelevant in equilibrium.

Our solution concept is Perfect Bayesian equilibrium in its standard definition. We analyse pure strategies.

3 Analysis

As we consider the compensation included in an apology, t , as exogenous, we impose the following restrictions on the parameter to focus on interesting cases²⁰.

1. Apologies work: if information were complete d_s will tender an apology to P_F , and P_F will accept it irrespective of the trial's outcome²¹.
2. Apologizing is risky, i.e. not every plaintiff would be ready to forgive. This was the motivation for legislators to introduce apology laws. In order to have P_L not willing to accept the compensation included in the apology we assume $t < \lambda M$: the compensation offered is lower than the minimum expected award at trial when D is guilty²².

The below assumption on t summarizes these conditions²³.

Assumption: $\max\{0; M - v_F\} < t < \lambda M$ (A)

We next introduce the following lemmas that show the two types of equilibria that can exist in both games. Lemma 1 shows the unique types of litigants that may settle in equilibrium. Lemma 2 establishes the pooling equilibria that exists when there is no settlement

²⁰We consider that this amount is not the result of a bargaining outcome between litigants and analyse the games for different values of t .

²¹Notice that we do not analyse the possibility that apology laws encourage settlement because courts become less efficient as less evidence is available to them. For this we restrict the analysis to $v_F + t > M$: P_F will accept an apology from d_s even when she expects the maximum amount of indemnity payments at trial

²²This condition also enables us to analyse the law's effect on insincere apologies. As $t < M$, d_n will prefer to apologize and pay t to P_F rather than pay the indemnity payments of the case, and P_F will prefer to not accept an apology from this type.

²³(A) implicitly assumes $\lambda > 1 - v_F/M$. The trial process has a prior positive probability to uncover the truth if $M > v_F$.

outcome in equilibrium. In order to simplify the exposition we present the analysis for the case where $c > 0$, but all the results hold for $c = 0$.

Lemma 1: (*Settlement types*) *Whenever there is a settlement outcome in equilibrium in both games, it must be between d_s and P_F .*

PROOF. d_i never settles as we have assumed she never apologizes. P_L neither as we have assumed $t < \lambda M$, so she prefers an expected award of at least λM at trial when receives an apology from a guilty type. d_n cannot settle alone, as if this is the only type that apologizes P_F will not accept it. Finally, we need to prove that there cannot be an equilibrium where both guilty types, d_s and d_n , apologize and the apologies are accepted by P_F . The key assumption in the argument is that there are no frivolous claims. Assume that we have an equilibrium where both guilty types settle. If it were the case, both plaintiffs will drop the case after receiving no apologies, as only d_i not apologizes. Given this, d_n will prefer to deviate to not apologize and pay nothing, as the case is dropped, instead of paying t to compensate P_F -or the cost of a rejected apology and subsequent trial with probability q_L . ■

Notice that the argument about the impossibility of having settlement with both guilty types relies on the assumption that both plaintiffs drop the case whenever the only d that not apologize is d_i . The result can also be proved by considering cases where P_F will not accept insincere apologies. To avoid the introduction of a new condition for this we rely on the above argument, but the reader can see this impossibility result as *apologies are non credible* that is more intuitive²⁴.

Lemma 2: (*Pooling equilibria*) Γ^a and Γ^{-a} always have equilibria which prescribes every D 's type to not apologize, and J awards damages $m = 0$ ($m = M$) if $p_i > (<) p_n + p_s$ when uninformed. These equilibria holds for P_F 's off-the-path equilibrium beliefs specifying that an apology must come from d_n .

²⁴Notice that P_F will reject apologies from both guilty types whenever $v_F \frac{p_s}{p_s+p_n} < T - t$, where T is the expected award at trial against a guilty type. The benefit from keeping the relationship with d_s is lower than the difference between the award and the settlement compensation with both types. When this hold the result in Lemma 1 applies.

PROOF. As no apology is tendered, P_F does not make a decision. Notice that given P_F 's off-the-path beliefs, she would reject an apology if tendered. Then, as apologies are costly, $c > 0$, and always rejected, guilty defendants will prefer to not apologize. As every defendant ends at trial, J 's award when uninformed will only depend on the prior evidence, $p_i > (<)p_n + p_s$. ■

We say that apology laws have an effect when there exists a difference in the settlement/trial equilibrium outcomes between games. By changing J 's available evidence, the introduction of apology laws will change the D 's incentives to apologize. In figure 1, panel a, we provide an overview of the results. An unintended effect of apology laws arises when players were playing an equilibrium with a settlement outcome before apology laws, that is broken after the introduction of apology laws. After apology laws d_n also prefers to apologize and by Lemma 1 this cannot be an equilibrium. In these cases players start playing the unique pooling equilibrium of Lemma 2 after apology laws. On the other hand, the intended effect arises when the opposite happens: from a pooling equilibrium to an equilibrium with settlement.

To understand the incentives of D , let us consider the trade-off involved in the decision of whether to apologize or not when apologies are admissible—in Γ^a . Notice that in equilibrium D will apologize if apologies are accepted by P_F . Otherwise, she will not apologize as apologies are costly when rejected ($c > 0$). In equilibrium the apology's decision will be given by the following inequality:

$$(1 - q_L)(v_d - t) - q_L(c + M) \geq -\lambda M - (1 - \lambda)m \quad (1)$$

The left-hand-side of (1) represents the payoff from apologizing. Whenever type d believes that P_F will accept an apology, with probability $(1 - q_L)$, she must compensate P_F by paying t and obtains v_d , the gains from restoring the relationship. However, she also faces the risk of encountering P_L with probability q_L . In this latter case P_L will take the apology to trial where then, as apologies are admissible and innocent types never apologize, she must pay $m = M$. Here she also incurs in the cost of a rejected apology, c . On the other hand, the right-hand-side represents the payoff from not apologizing where she will face a trial with an expected cost of $\lambda M + (1 - \lambda)m$: the trial uncovers the liability with probability λ , where she must pay M , and with probability $1 - \lambda$ J awards $m \in \{0, M\}$

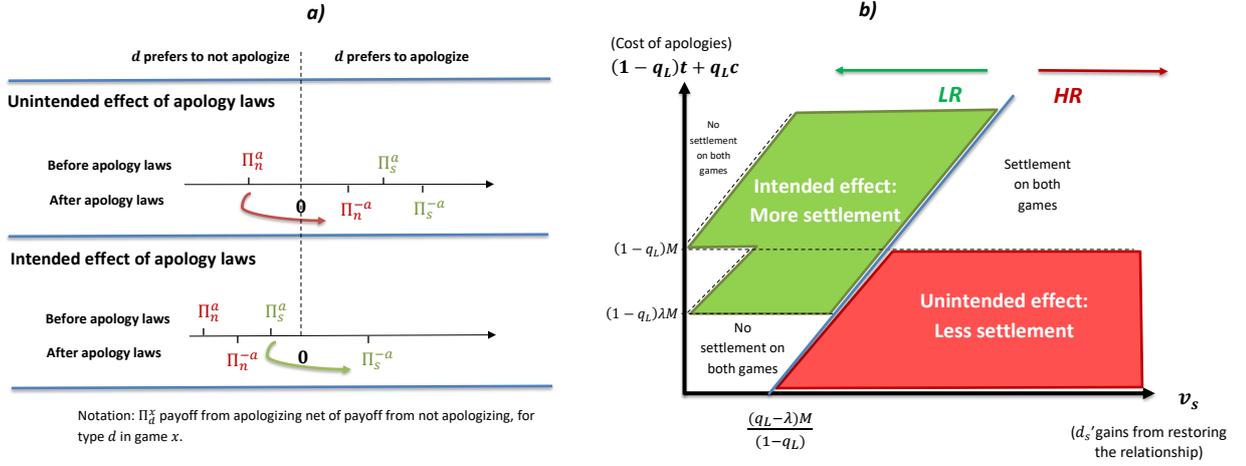


Figure 1: The effects of apology laws. a) How incentives to apologize change. b) Types of relationships affected (the graph is for a fix q_L and $q_L > \lambda$)

depending on her inference about which types end at trial without apologies.

Now we label an important condition that will be determinant for the identification of the effects of apology laws, and how apology laws affects different types of relationships.

Definition: We will say that d_s has *high gains from restoring the relationship (HR)* with P_F (resp. *low gains (LR)*) when the following condition is satisfied (resp. not satisfied)

$$(1 - q_L)(v_s - t) - q_L(c + M) > -\lambda M \quad \text{or equivalently,}$$

$$v_s > \frac{(q_L - \lambda)M + q_Lc}{1 - q_L} + t$$

That is, whenever d_s has *HR* she prefers to apologize and face the risk of paying M at trial when she encounters P_L instead of not doing it and pay λM . Notice that when d_s has *HR* she prefers to apologize despite that the award at trial when no apology is presented is the lowest that she can expect to pay, λM . On the other hand, whenever d_s has *LR* she prefers to remain silent in this scenario, apologies are too costly. In figure 1, panel b, we show that the effects of apology laws differs on both sides of this condition.

We provide the complete characterization of the equilibria of each game in the appendix while in the main text we focus on the space of parameters that delivers a difference in terms of settlement/trial outcomes between games. In the next subsections we explain the effects of apology laws in detail.

3.1 The unintended effect of apology laws

We start by showing an equilibrium with settlement when apologies are admitted, and then show how the introduction of apology laws breaks the settlement.

Lemma 3: (*Settlement before apology laws*) Γ^a has an equilibrium which prescribes type d_s alone to apologize, type P_F to accept an apology, and J to award damages of 0 at trial when no apology is presented and M when presented, if and only if

- i) d_s has HR ,
- ii) $(1 - q_L)t + q_L(c + M) > \lambda M$, and
- iii) $p_i > p_n$

PROOF. (Sufficiency) (*Plaintiff*) By construction, P and J would infer D 's type from an apology; so type P_F cannot profitably deviate from accepting. P_L always rejects an apology by assumption.

(*Judge*) By construction, the case would only go to trial without apologies if D 's type is either d_i or d_n . Hence, J would award 0 at trial when no apologies is presented – and she remains uninformed –, given $p_i > p_n$. On the other hand, when she observes an apology at trial, she would award $m = M$, as just d_s apologizes.

(*Defendant*) Consequently, d_s and d_n lose λM by not apologizing. The case must then proceed to trial without apologies where they pay 0 with probability $1 - \lambda$, i.e. whenever J remains uninformed. Then, given that P_L would always turn down an apology, we have

$$d_s \text{'s apologies payoff} : (1 - q_L)(v_s - t) - q_L(c + M) > -\lambda M$$

$$d_n \text{'s apologies payoff} : (1 - q_L)(-t) - q_L(c + M) < -\lambda M$$

The first inequality holds as d_s has *HR*; thus d_s cannot profitably deviate to remain silent. The second inequality follows by ii, and d_n cannot profitably deviate to apologize. Finally, d_i never apologizes by assumption.

(Necessity) Suppose d_s has *LR* (condition i does not hold). Then d_s can profitably deviate to not apologize as the cost of apologizing, $(1 - q_L)(v_s - t) - q_L(M + c)$, exceeds $-\lambda M$, the expected award she must pay at trial without apologies. This proves necessity of condition i. In the same sense, if ii does not hold, then d_n would prefer to deviate to apologize. Finally, assume the prescribed equilibrium with $p_n > p_i$. By our prescribed equilibrium, the types that end at trial without apologies presented are d_i and d_n . Hence, now J would award M at trial when uninformed and no apology is presented. Equilibrium a2 in the appendix considers this case. ■

Notice the features that allow settlement in this equilibrium: 1) d_s has *HR*: she prefers to take the risk of apologizing despite apologies are admissible, and 2) high enough p_i (condition iii): d_n prefers to remain silent to be pooled with d_i on trial.

Consider now that apology laws are introduced in an environment where conditions *i – iii* of the above equilibrium hold. We will show that apology laws induce insincere apologies in these environments for two different mechanisms. First, because apology laws reduce the legal consequences of apologies as they cannot be presented as evidence, *cheaper apologies*. Second, because of *J's higher awards*: as now J cannot observe sincere apologies when they are tendered and rejected by P_L in trial, she awards indemnity payments more often than before because of an adverse inference about liability. Both effects cause the previous settlement outcome to fail by Lemma 1.

Proposition 1: *(The unintended effect) The introduction of apology laws reduces settlement between litigants whenever the equilibrium with settlement before apology laws exist (Lemma 3) and*

- a) *(Cheaper apologies)* $(1 - q_L)t + q_Lc < (1 - q_L)\lambda M$ and $p_i > p_n + q_L p_s$, or
- b) *(J's higher awards)* $(1 - q_L)t + q_Lc < (1 - q_L)M$ and $p_i < p_n + q_L p_s$

PROOF. (a) *Cheaper apologies.*) Towards a contradiction, assume that we have the equilibrium strategies for P and D of Lemma 3 after the introduction of apology laws. Notice that given these strategies and $p_i > p_n + q_L p_s$, J would award $m = 0$ at trial when uninformed. d_i and d_n end at trial because they do not apologize, and d_s because she apologizes and P_L rejects it. As J cannot distinguish between cases that reach trial with and without apologies, $p_i > p_n + q_L p_s$ implies her best reply of $m = 0$. Consider then d_n 's incentives to apologize. Notice that now d_n faces a lower risk from apologizing: the new cost from apologizing is $(1 - q_L)t + q_L(c + \lambda M)$; lower than before apology laws, $(1 - q_L)t + q_L(c + M)$. This arises as apologies cannot be used by P_L as evidence after the introduction of apology laws and J awards $m = 0$ in the unique trial outcome. Then, if $(1 - q_L)t + q_L(c + \lambda M) < \lambda M$ d_n will prefer to apologize, as this cost is lower than the cost of remaining silent, λM . But we have already proved in Lemma 1 that there cannot be settlement with both guilty types.

(b) *J's higher awards.*) Towards a contradiction, assume that we have the equilibrium strategies for P and D of Lemma 3 after the introduction of apology laws. Notice that given these strategies and $p_i < p_n + q_L p_s$, J would award $m = M$ at trial when uninformed. d_i and d_n end at trial because they do not apologize, and d_s because she apologizes and P_L rejects it. As J cannot distinguish between cases that reach trial with and without apologies, $p_i < p_n + q_L p_s$ implies her best reply of $m = M$. Consider then d_n 's incentives to apologize. As $(1 - q_L)t + q_L(c + M) < M$ d_n will prefer to apologize whenever the award at trial is M ; she avoids to pay M by apologizing as P_F accepts it in our prescribe equilibrium. The cost of apologizing is lower than the cost of remaining silent, M . But we have already proved in Lemma 1 that there cannot be settlement with both guilty types. ■

3.1.1 Relationships affected by apology laws

Apology laws have effects on the type of relationships than can be restored. Figure 1, panel b, shows the effects of apology laws as a function of d_s 's incentives to restore the relationship. Notice that overall apology laws decrease settlement whenever d_s has high gains from restoring the relationship with P_F , HR .

Let us define the *value of a restored relationship* by $v_F + v_s$. The higher this amount, the higher the social value of rules that promotes the recovery of the relationship. Recall from (A) that t may depend negatively on v_F . The higher the value P_F obtains by recovering the relationship with d_s , v_F , the lower the compensation needed to restore it, t . In this sense,

notice from Figure 1, panel b, that apology laws start working when the cost of apologies is high enough, and do not work whenever this cost is low ($((1 - q_L)t + q_L c < -(1 - q_L)\lambda M$). The following corollary follows from this and propositions 2,

Corollary 1: *(Relationships affected) apology laws are detrimental to restore socially–highly valuable relationships.*

3.2 Miscarriages of justice

In this section we compare the miscarriages of justice for each game where the unintended effect holds. For the analysis we extend the payoffs of J in line with previous literature (see Feddersen and Pesendorfer (1998)). We assume that there is $\delta \in (0, 1)$ such that the jury earns a pay-off of: 0 whenever the award is correct, δ whenever an innocent defendant must pay, and $(1 - \delta)$ whenever it acquits a liable defendant. δ is the standard of proof: the critical probability of guilt at which the jury is indifferent between acquitting and convicting at trial. The analysis in previous section has implicitly assumed $\delta = 0.5$ ²⁵.

Let us first consider the miscarriages of justice on the equilibrium before apology laws—the one described in Lemma 3. d_s settles with P_F and pays damages $m = M$ to P_L as the latter presents the apologies in court. On the other hand, d_i and d_n end at trial without apologies were then they do not pay indemnity payments, $m = 0$, when J is uninformed. So the miscarriages of justice is the acquittal of d_n to pay damages.

Recall that when the unintended effect arises after the introduction of apology laws players start playing one of the pooling equilibrium described in Lemma 2. Which equilibrium is played depends on whether $\delta \geq \frac{p_n + p_s}{p_i + p_n + p_s (=1)}$ (before $p_i \geq p_n + p_s$, see Lemma 2).

1. If $\delta > \frac{p_n + p_s}{p_i + p_n + p_s (=1)}$, then $m = 0$ at trial after apology laws. In this cases, p_n and p_s are acquitted from paying damages. So after apology laws the miscarriages of justice increases by p_s .

²⁵ $\delta = 0.5$ is a reasonable assumption. Notice that the standard of proof in a civil case is proof on the balance of probabilities, and this means that the party bearing the burden of proof must prove that her case is more probable than not.

2. If $\delta < p_n + p_s$, then $m = M$ at trial after apology laws. Here p_i must pay damages, but also guilty types. To compare the miscarriages of justice between games we need to consider whether the acquittal of d_n is preferred to make d_i paying, $-(1 - \delta)p_n \geq -\delta p_i$. So in order to have higher payoff for J under apology laws we need $\delta < p_n / (p_i + p_n)$. Notice that we are considering cases where $\delta p_i > (1 - \delta)p_n$ (condition iii of Lemma 3), so if $\delta = 0.5$ apology laws always incur in higher miscarriages of justice.

Corollary 2: (*Miscarriages of justice*) *Apology laws incur in higher miscarriages of justice whenever the unintended effect holds (proposition 1) unless the standard of proof, δ , is low enough ($\delta < \min\{p_n + p_s, \frac{p_n}{p_i + p_n}\}$). In particular, if $\delta = 0.5$ apology laws always incur in higher miscarriages of justice.*

[McMichael et al., 2019] shows that apology laws increase indemnity payments for non-surgeons. Our results predict that this is because more innocent physicians end at trial (case $\delta < p_n + p_s$).

3.3 The intended effect of apology laws

In this section we show that apologies may increase settlement, that is, the intended effect. Then we show that for some cases they achieve settlement but at the cost of affecting the reliability of courts.

The reasons for why apology laws may increase settlement are analogous to the ones provided in the previous section, but now the effects work through d_s 's incentives to apologize rather than on d_n 's ones. First, *cheaper apologies*: they may increase settlement by reducing the legal consequences of apologies for sincere defendants, that refrain to apologize if apologies were admissible. This is what motivates legislators to introduce apology laws in the first place. Second, *J's higher awards*: they may increase settlement by increasing the cost of not apologizing. J 's higher awards under apology laws encourage apologies from d_s more often than before these laws. Importantly, for the intended effect to arise these mechanism must encourage apologies from d_s but not from d_n ; so the mechanism must not be too strong to encourage d_n apologies. This can be seen in Figure 1, panel b,

where the intended effect holds for high enough cost of apologizing. The following proposition shows these results, that we prove it in the appendix.

Proposition 2: *(The intended effect) Apology laws induce more settlement by reducing the legal cost of apologies whenever d_s has LR and*

- a) *(Cheaper apologies) $v_s > t - \lambda M + \frac{q_L c}{1 - q_L} > 0$, $p_i > p_n + q_L p_s$; or*
- b) *(J's higher awards) $v_s > t - M + \frac{q_L c}{1 - q_L} > 0$, $p_n < p_i < p_n + q_L p_s$; or*
- c) *(No insincere D) $p_n = 0$ and $v_s > t - \lambda M + \frac{q_L c}{1 - q_L}$*

Notice that in *cheaper apologies* these laws achieve settlement by excluding the possibility to use apologies at court by P_L . In the equilibrium after apology laws P_L has evidence that proves the liability of d_s in equilibrium, i.e. d_s 's apologies. However, she cannot use it at trial, where then d_s is exonerated to pay damages. This highlights one of the arguments raised against apology laws: Public confidence in the courts could be adversely affected if a person who has admitted liability in an apology is found not liable (Cohen [2002]).

Corollary 3: *Apology laws may induce settlement at the cost of affecting the public confidence in courts.*

Consider the role play by c , the cost of a rejected apologies. Chaudhry and Loewenstein [2019] proposes that apologies are effective because they involve giving up something valuable: being perceived as competent. Transgressors who apologize in situations in which competence is relevant suffer a negative impact on their perceived competence ([Kim et al., 2006]; [Weiner et al., 1991]), and speakers are aware of this. In this sense, c is related to competence and its value will be dependant of the type of relationship considered. Physicians are reluctant to give apologies because it damages their reputation (Gallagher et al. [2003]). Notice that whenever $c \sim 0$, as we assume $t < \lambda M$, apologies cannot work, as $(1 - q_L)t < (1 - q_L)\lambda M$ (see Figure 1, panel b),

Corollary 4: *Apology laws can only work on relationships where competence is damaged by apologizing.*

3.4 Frivolous claims and lower settlement amounts

As shown in Lemma 1, if there are no frivolous claims we cannot have an equilibrium where both guilty types settle in any game. Now we show that when frivolous claims are considered, apology laws may facilitate settlement through insincere apologies. That is, cases that before apology laws feature settlement just with d_s , after the introduction of these laws feature settlement with d_s and d_n . This result highlights one of the arguments made by apology laws detractors: apology laws will encourage the manipulation of naive injured parties (not to pursue money damages, not to speak to a lawyer and not to become educated about their legal rights) by sophisticated actors (insurers, hospitals, big companies) (Cohen [2002])²⁶.

Usually attorneys in civil claims receive a contingent fee, where the attorney receives a percentage of any settlement or judgement but receives nothing if the case is lost. Moreover, sometimes attorneys have the control over the settlement decision (see Spier [2007] for a discussion). We model this by assuming that some proportion of plaintiffs, $x \in (0, 1)$, have this type of contracts and this is common knowledge by players. Their attorney will take innocent types to trial if she expect positive payments. Let us denote the probability of facing a plaintiff with this type of contract as $x(q_F + q_L) = x$. Notice that not all litigious plaintiffs, P_L (the ones that do not accept apologies), will take innocent types to trial provided they know the defendant is innocent, just a fraction xq_L .

Denote the expected awards at trial for guilty types as $T \in \{\lambda M, M\}$, and assume $c = 0$ for simplicity.

Proposition 3: *(Forgiving insincere defendants) Apology laws encourage settlement through the acceptance of insincere apologies whenever:*

$$\begin{aligned} i) \quad & x > \frac{(1-q_L)t + q_L(c+T)}{(1-q_L)t + T} \\ ii) \quad & v_F \frac{p_s}{p_s + p_n} > T - t. \end{aligned}$$

²⁶Proposition 3 shows that some plaintiffs are accepting apologies from insincere defendants with positive probability in equilibrium under apology laws. If we consider that the compensation included in the apology in these cases, t , is similar to the one included when just a sincere defendants apologize—before apology laws, then we show another argument against apology laws: “Apologies encouraged by such legislation might create an emotional vulnerability in some plaintiffs who may accept settlements that are inappropriately low.” (British Columbia Ministry of Attorney General (2006), p.4).

PROOF. We show first that we cannot have an equilibrium with this feature in Γ^a . Assume we have an equilibrium in which both guilty types settle. Notice that if both types apologize, if $m = 0$ when no apologies are presented at trial, frivolous claims against innocent types will not be filled, as attorneys do not expect gains. Then, d_n will prefer to mimic innocent defendants and pay nothing as in the benchmark model.

On the contrary, in Γ^{-a} , the impossibility for J to distinguish between cases with and without apologies makes frivolous claims possible. Now, by mimicking d_i guilty types could end in trial through frivolous claims. Assume we have an equilibrium were both guilty types apologize and settle, and J awards m . J awards at trial will depend on whether the probability of frivolous claims is higher or lower than cases with rejected apologies, if $x p_i \leq (q_L + x q_F)(p_s + p_n)$. Then, if

$$v_F p_s + t(p_n + p_s) > (\lambda M + (1 - \lambda)m)(p_s + p_n) = T(p_s + p_n)$$

P_F will accept apologies for both guilty types. Finally d_n will compare the payoffs from apologizing with the ones from not doing it, that depends on the trial award:

$$(1 - q_L)(1 - x)t + q_L(c + T) < xT$$

Hence, whenever the above inequality holds, d_n will prefer to apologize. If type d_n settles in equilibrium, it must be that type d_s also does. We have an equilibrium. ■

4 Discussion

Apologies from innocent types

In our model an apology is an admission of fault; when accepted it implies a compensation. To analyse apologies form innocent types we should differentiate two interpretations of

apologies: excuses vs. admissions. Outside the legal context, [Ho, 2012] introduces just two types of agents, high and low in terms of future gains for the principal, and do not model the liability of the agent: when the output for the principal is low it may be because of external circumstances or because of an agent's low effort. Then his model cannot differentiate between excuses and admissions. In his setting, in any separating equilibrium just the high type apologize. Then, we can interpret this apologies as excuses, if consider high types as innocent ones; or admissions, if a high type is guilty but with high gains from keeping the relationship. As our model differentiate the liability of the agents, innocent or guilty, from the payoffs for future gains, i.e. sincerity, we can analyse this distinction. More importantly, this distinction is necessary to explain the unintended effect of apologies laws. Below we show how our results extend to apologies for d_i in each interpretation:

1. Excuses. Notice that our equilibrium in Lemma 3 can be reinterpreted as one in which d_i and d_n offer costless excuses, and d_s confess. If we assume that no P will accept excuses then we are in the same situation as in the benchmark interpretation. If we assume that some excuses are accepted by P , and so cases are dropped without compensation, then we would need higher v_s in order for d_s to apologize. But this maintain the intuition of our results.
2. Admissions. In some context innocent types may also prefer to take responsibility for what happen if they can avoid a negative sentence at trial. Two factors allow the strategies of Lemma 3 to be robust to this: i) d_i would have to pay a lower amount at trial whenever $\lambda > 0$, so her incentives to not confess would generally be higher; and ii) false confession are insincere apologies, tendered to avoid trial. Notice that our history starts after the relationship is broken, P will take the case to trial if no apologies are tendered. So even if we assume that d_i has $v_i > 0$, i.e. she also obtains a benefit from restoring the relationship with P_F , it seems natural to think that $v_i < v_s$. In the context of medical malpractice, the assumption that innocent physicians are reluctant to provide false confession has some support²⁷.

²⁷A survey when a ransom sample of 3.985 physicians were asked if saying "I'm sorry" would have helped, the answer was 'No' in 81% of responses. "Among the verbal comments to this question, most physicians reported that they didn't say they were sorry because it wasn't their fault...". This results remains for a similar survey in years 2017 and 2019. (<https://www.medscape.com/features/slideshow/public/malpractice-report-2015page=25>, accessed on 25/09/2020). Moreover, Golann [2011] shows that most claims are dropped by

Comments on McMichael et al. [2019]

1. Comparing different mechanism. [McMichael et al., 2019] explain their results by an hypothesis that combines i) more apologies from physicians after apology laws, and ii) apologies as signal of malpractice. Their argument is that after apology laws physicians “rush to apologize following an error without completely understanding the risks and complexities of apologizing in the wake of an error”²⁸. The difference in outcomes between surgeons and non-surgeons is explained by considering apologies as signal of malpractice. Asymmetric information about liability, i.e. whether the plaintiff knows the defendant is guilty or not, is more likely to be present in malpractice claims involving non-surgeons than surgeons²⁹. Then, if physicians apologies more often after apology laws, previous malpractice cases that were not filled because were not noticed by patients, mostly for non-surgeons ones, would end in court more often. Our results offer an alternative explanation based on the physicians’ valuation of the relationship.

One way to test which mechanism is present would be to consider whether the unintended effect of apology laws decrease over time or is stable³⁰. As [McMichael et al., 2019] argue “If apology laws increase malpractice risk and, on balance, are not in their best interests, why would physicians continue to apologize? While future research should investigate this question in detail, the most likely answer is that physicians have simply been conditioned to apologize with little training on how to do so effectively...Moreover, many physicians are not involved in multiple malpractice actions, so they have little reason to know—particularly given the positive treatment of apologies from a variety of sources—that apologizing can increase their malpractice liability risk”³¹. One could expect

the plaintiff after filled (%58): as plaintiffs acquire more information in the course of a lawsuit, they often conclude that a claim is weaker than they had first thought. This should deter false admissions.

²⁸[McMichael et al., 2019], p.390.

²⁹[McMichael et al., 2019], p.369: “Surgeons generally interact with and treat patients in a discrete event, i.e., the surgery they are performing plus any pre-operative and post-operative care. Because of this discrete interaction, patients who suffer an injury will likely have little trouble tracing that injury to an error that occurred during surgery. On the other hand, non-surgeons generally treat their patients over the course of years or may interact with patients a number of times when attempting to resolve an injury or illness. Thus, observing the malpractice of non-surgeon physicians may be more difficult”.

³⁰This can be easily done in the empirical analysis of [McMichael et al., 2019] but unfortunately the data is not public available

³¹[McMichael et al., 2019], p.389.

that this hypothesis, if holds, has a short run impact. Once physicians learn that apologies increase malpractice risk, no more apologies would be tendered. If apologies as signal of malpractice is the responsible for the unintended effect, one should expect that the effect will diminish over time. On the other hand, our theory, as an equilibrium result, predicts the long term effect of apology laws once player learn how to play. The unintended effect should be more stable.

2. Policy implication. The hypothesis of more physicians' apologies after apology laws makes [McMichael et al., 2019] to moderate the advice of repealing these laws³². "To the extent that apology laws promote transparency in the physician-patient relationship through the revelation of otherwise hidden malpractice, they may benefit society. Indeed, this transparency may elucidate errors that would have been repeated but for the apology that was offered. Because this increased transparency comes at the cost of increased malpractice liability risk, state lawmakers must weigh transparency against liability in deciding whether to repeal apology laws"³³. Our results predict that not necessarily more transparency is achieved after apology laws. If there is a small cost to apologize for physicians when they know that an apology will not be rejected ($c = \varepsilon > 0$), apology laws also reduce communication in our model.

³²[McMichael et al., 2019], p.390:"The most natural course of action may be to repeal these laws, given their specific inability to achieve their stated purpose....If apology laws—even unintentionally—promote apologies that improve people's lives, they may generate a net social benefit with respect to patients' well-being, despite their failure to achieve their primary financial goal".

³³[McMichael et al., 2019], p.391.

Appendix

Characterization of equilibria in each game

Lemma A: *Settlement in the admissible case*

a1) Γ^a has an equilibrium which prescribes type d_s alone to apologize, type P_F to accept an apology, and J to award damages of 0 at trial if no apology is presented and M otherwise when uninformed, if and only if

i) $(1 - q_L)(v_s - t) - q_L c > (q_L - \lambda)M$,

ii) $(q_L - \lambda)M > -(1 - q_L)t - q_L c$, and

iii) $p_i > p_n$

a2) Γ^a has an equilibrium which prescribes type d_s alone to apologize, type P_F to accept an apology, and J to award damages of M at trial irrespective of the presentation of an apology when uninformed, if and only if

i) $(1 - q_L)(v_s - t) - q_L c > -(1 - q_L)M$,

ii) $-(1 - q_L)M > -(1 - q_L)t - q_L c$, and

iii) $p_i < p_n$

These are the unique equilibria of Γ^a

PROOF: (Uniqueness) We just need to consider equilibria that depends on m when no apology is presented: $m = 0$ (equilibrium a1) and $m = M$ (equilibrium a2). Equilibrium a1 was proved in Lemma 3.

Proof of equilibrium a2): (Sufficiency) By construction, P would infer D 's type from an apology; so type P_F cannot profitably deviate from accepting.

Moreover, by construction, the case would only go to trial without an apology if D 's type is either d_i or d_n . Hence, as $p_n > p_i$, J would award M at trial when no apology is presented and she remains uninformed. On the other hand, she would award $m = M$ when observes an apology, as only d_s apologizes.

Consequently, type d_n loses M by not apologizing and $(1 - q_L)t + q_L(M + c)$ by apologizing (as P_L never accepts an apology); hence, $(1 - q_L)t + q_L(M + c) > M$ implies that she cannot profitably deviate. Type d_s cannot profitably deviate neither: she loses $-(1 - q_L)(v_s - t) + q_L(M + c)$ by apologizing, lower than M , the payoff from not doing it, given condition i.

(Necessity) Suppose that $(1 - q_L)(v_s - t) - q_L(M + c) < -M$ (condition i does not hold). Then, type d_s can deviate to not apologize as the cost of apologizing exceeds $-M$, the expected award she must pay at trial without apologies. This proves necessity of condition i. In the same sense, if ii does not hold, then d_n would prefer to deviate to apologize. Finally, assume the prescribed equilibrium with $p_n < p_i$ (condition iii does not hold). By our prescribed equilibrium, the types that end at trial without apologies presented are d_i and d_n , hence, J would award 0 at trial when uninformed. Equilibrium a1 already considered this case. ■

Lemma B: *Settlement in the inadmissible case*

a1) Γ^{-a} has an equilibrium which prescribes type d_s alone to apologize, type P_F to accept an apology, and J to award damages of 0 at trial when uninformed, if and only if

- i) $(1 - q_L)(v_s - t) - q_Lc > -(1 - q_L)\lambda M$,*
- ii) $-(1 - q_L)\lambda M > -(1 - q_L)t - q_Lc$, and*
- iii) $p_i > p_n + q_Lp_s$*

a2) Γ^{-a} has an equilibrium which prescribes type d_s alone to apologize, type P_F to accept an apology, and J to award damages of M at trial when uninformed, if and only if

- i) $(1 - q_L)(v_s - t) - q_Lc > -(1 - q_L)M$,*
- ii) $-(1 - q_L)M > -(1 - q_L)t - q_Lc$, and*
- iii) $p_i < p_n + q_Lp_s$*

These are the unique equilibria of Γ^{-a}

PROOF: (Uniqueness) We just need to consider equilibria that depends on m : $m = 0$ (equilibrium a1) and $m = M$ (equilibrium a2).

Let us start with a1. By construction, P would infer D 's type from an apology; so type P_F cannot profitably deviate from accepting.

The cases that go to trial are between P_L and every D 's type; and between P_F if the D is d_n or d_i , as the they do not apologize. Hence, J would award 0 when uninformed at trial because $p_i > p_n + q_L p_s$ (this proves necessity of $p_i > p_n + q_L p_s$).

Consequently, types d_n and d_s would respectively lose $(1 - q_L)t + q_L(\lambda M + c)$ and $(1 - q_L)(t - v_s) + q_L(\lambda M + c)$ if they apologize. As $m = 0$ when J is uninformed, type d_s prefers to apologize given $(1 - q_L)(v_s - t) - q_L c > -(1 - q_L)\lambda M$ (this proves necessity of condition i). In the same sense, type d_n prefers to not apologize as $-(1 - q_L)\lambda M > -(1 - q_L)t - q_L c$.

Consider now a2). By construction, P would infer D 's type from an apology; so type P_F cannot profitably deviate from accepting. The cases that go to trial are between P_L and every D 's type; and between P_F if the D is d_n or d_i , as the they do not apologize. Hence, J would award M when uninformed at trial because $p_i < p_n + q_L p_s$ (this proves necessity of $p_i < p_n + q_L p_s$).

Consequently, as F forgives, types d_n and d_s would respectively lose $(1 - q_L)t + q_L(M + c)$ and $(1 - q_L)(t - v_s) + q_L(M + c)$ if they apologize. Type d_n prefers to not apologize as by apologizing she loses a higher amount than M , the cost of not doing it (this proves necessity of condition ii). Type d_s prefers to apologize as $(1 - q_L)(v_s - t) - q_L(c + M) > -M$ (this proves necessity of condition i). ■

Proof of Proposition 2

Based on the above characterization, let us start with part a) of proposition 2, cheaper apologies. Consider the equilibrium a1 of Γ^{-a} . Condition iii precludes the existence of equilibrium a2 in Γ^a (as $p_i > p_n$). On the other hand, if d_s has LR equilibrium a1 of Γ^a cannot exist. Then whenever equilibrium a1 of Γ^{-a} exists and d_s has LR we have an equilibrium with settlement under apology laws, and just the pooling equilibria of Lemma 2 when apologies are admissible.

For part b), J 's higher awards, consider the equilibrium a2 of Γ^{-a} . It has the same existence conditions i-ii as equilibrium a2 in Γ^a , but holds for more cases regarding condition iii;

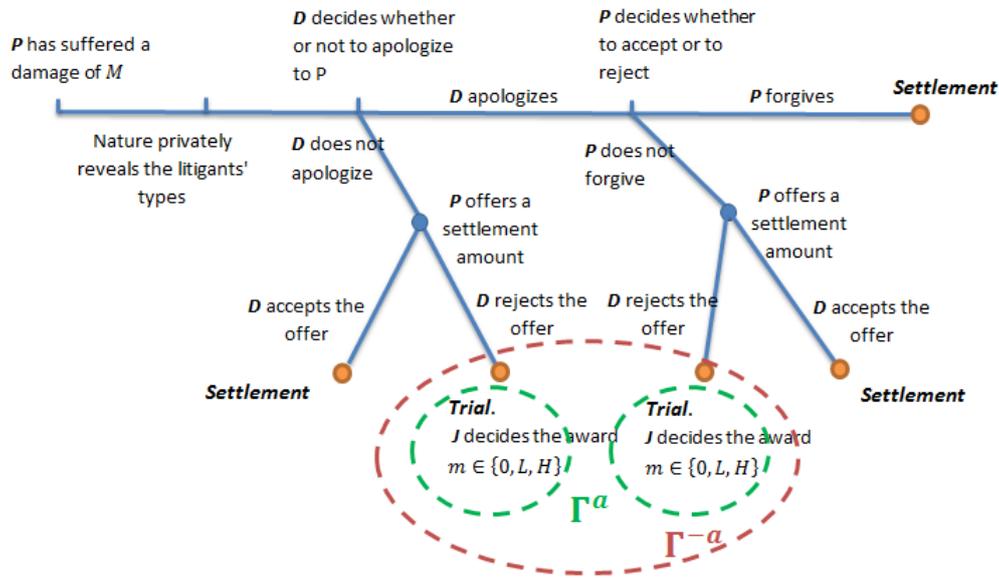
whenever $p_i > p_n$. Then whenever equilibrium a2 of Γ^{-a} exists, $p_i > p_n$, and d_s has LR , we have an equilibrium with settlement under apology laws, and just the pooling equilibria of Lemma 2 when apologies are admissible.

Finally, part c. If $d_n = 0$, an apology will always be accepted by P_F as there are just sincere defendants. We just need to consider the incentives of d_s to apologize that depends on the cost of apologies. Then, as apologies are cheaper under Γ^{-a} , d_s will apologize for more cases in the inadmissible case. ■

Settlement bargaining and offers

In this section we extend the benchmark model to incorporate settlement bargaining. We consider the most simple approach, an ultimatum screening game a la Bebchuk [1984]³⁴. After an apology's rejection, or in the case where D chooses to not apologize, P can offer a settlement amount to D denoted by $z(a)$. Then, if D accepts the offer the case settles, otherwise goes to trial. The next figure shows the extended timeline. Notice that the offer will be conditional on the apologies decision, $a \in \{0, 1\}$, where $a = 1(0)$ denotes that an apology was (not) tendered. Hence, after the apologies decision the extensive form is equal to Bebchuk [1984] screening model. We consider three types for liability: innocent, low liability and high liability. J 's awards m belong to $\{0, L, H\}$, where $L, H > 0$ implies that D was found liable in the case, and $L, H > 0$ implies that D was found liable in the case, and L denotes a lower level of liability respect of H , $L < H$. We now have five types of D : innocent (i), guilty—sincere with low (sl) and high (sh) liability, and finally, guilty—insincere with each level of liability in the case, nl, nh . The prior probabilities are denoted $p_i, p_{sl}, p_{sh}, p_{nl}, p_{nh}$.

³⁴Despite its simplicity, the predictions of the ultimatum game are robust to a more general sequence of offers. See Daughety and Reinganum [2012] survey about settlement bargaining



We analyse the case where innocents' D do not apologies. Notice that if apologies are not tendered, then P_F will have the same payoff function as P_L . This implies that both types will offer the same settlement amount if apologies are not tendered, or rejected by P_F . The only difference between P 's types arises in case apologies are offered: P_F would like to forgive sincere apologies while P_L does not. To avoid the introduction of trial cost for every player, we assume that when indifferent between trial and settlement, players prefer settlement. Moreover, as we do not introduce trial cost we need the following assumptions to ensure that screening types for the plaintiff is at least as good as going to trial with every type: $p_{gl} > p_{gh}$ for $g \in \{s, n\}$. Notice that now we have two types of settlement. First, the settlement driven by an accepted apology, that we call it settlement induced by apologies. Second, the settlement that results from bargaining, below we refer to this type of settlement unless otherwise specified. For the former we set the settlement amount by t , as in the benchmark model. For the latter, settlement amounts are equilibrium objects, and this is the amount that we are interest in. The following lemma will be useful for the analysis, it is the extension of the arguments of Bebchuk [1984] for three types and J imperfectly informed:

Lemma: *The screening settlement bargaining game implies that cases settles with the type with the highest liability involved in bargaining and goes to trial with the other(s). In*

particular, when we have three types of liability, i, l, h , cases settle with h -types and go to trial with l and i . Whenever P faces just two types, cases settle with the type with higher liability and go to trial with the other.

PROOF: Consider first the case where P faces the three types, i, l, h . Denote their respective probabilities, q_j for $j \in \{i, l, h\}$. P has the following options: i) offer z_p that pools every type, i.e. every type accept this offer; ii) offer z_l that is accepted by both guilty types and rejected by i ; iii) offer z_h that is accepted just by type h ; and iv) offer z_t that is rejected by every type.

Then notice that as we assumed $q_l > q_h$, when uninformed, J will award either L or 0 . Denote the award set by J when uninformed by X_j for $j \in \{p, l, h, t\}$ that refers to the indexes for the above settlement offers. Then we prove that z_h is the optimal offer. Notice that as z_h is accepted just by h , the best offer is to leave her indifferent, hence $z_h = \lambda H + (1 - \lambda)X_h$. Then the payoff from using z_h are:

$$\begin{aligned} \Pi(z_h) &= q_i(1 - \lambda)X_h + q_l(\lambda L + (1 - \lambda)X_h) + q_h(z_h) = \\ &= q_i(1 - \lambda)X_h + q_l(\lambda L + (1 - \lambda)X_h) + q_h(\lambda H + (1 - \lambda)X_h) = \\ &= (1 - \lambda)X_h + q_l\lambda L + q_h\lambda H \end{aligned}$$

Let us first compare the payoff from offering z_h and z_p . Notice that $z_p = 0$ for off-the-path beliefs of J specifying that the case that fails to settle must come from an innocent defendant. Hence if guilty types deviates they will pay for $m \in \{L, H\}$, and i is indifferent. Other off-the-path belief of J that implies higher z_p would not survive the intuitive criteria as the only type that is indifferent from a pooling offer is always the innocent type. Then $z_h > 0 = z_p$.

Now notice that $z_l = \lambda L$. By definition both type accepts z_l , then it must be that $X_l = 0$. Given this, l is indifferent between accepting or rejecting z_l , while h strictly prefers to settle. Then we show that the payoffs from z_h are higher than using z_l

$$\Pi(z_h) = (1 - \lambda)X_h + q_l\lambda L + q_h\lambda H > (q_l + q_h)\lambda L = q_i(1 - \lambda)X_l + (q_l + q_h)(z_l) = \Pi(z_l)$$

Finally, we show that $z_h = z_l$. Notice that as $q_l > q_h$ when P offers z_l , that every type rejects it by definition, $X_t \in \{0, L\}$, and will depend on whether $q_i \leq q_l$. Hence $X_t = X_h$. Then,

$$\begin{aligned} \Pi(z_h) &= (1 - \lambda)X_h + q_l\lambda L + q_h\lambda H = (1 - \lambda)X_t + q_l\lambda L + q_h\lambda H = \\ &= q_i(1 - \lambda)X_t + q_l(\lambda L + (1 - \lambda)X_t) + q_h(\lambda H + (1 - \lambda)X_t) = \Pi(z_l) \end{aligned}$$

With an analogous proof we can show that whenever P faces types l and h , she offers a settlement amount that leaves indifferent type h whereas l rejects the offer. The same holds for bargaining with $\{i, h\}$ or $\{i, l\}$. ■

Now we show that the results in section 3 hold in this extensive form. Apology laws reduce the settlement induced by apologies by encouraging insincere apologies. To see this consider the conditions for existence of equilibrium a1 in Γ^a , that do not have a counterpart in Γ^{-a} . This equilibrium is the analogous of the one shown in Lemma 3 in the benchmark model.

As this extension allows us to analyse the bargaining outcomes, we focus on the size of the settlement offers. We show that apology laws weakened the bargaining position of the plaintiff, and they settle for lower amounts. The reason is that plaintiffs that are not willing to forgive, before apology laws were introduced use the fact that an apology was tendered to ask for a higher settlement amount. After apology laws, given that apologies cannot be used as a threat any more, they settle for a lower amount. Moreover, the settlement amount is the same whether an apology is tendered or not; despite knowing that the defendant is guilty in the former case.

Proposition 4: *For cases that settles before and after apology laws, after the introduction of apology laws parties settle for lower amounts.*

PROOF: We first characterize the equilibria of each game as in the benchmark model. It would be useful in order to save space to rewrite

$$(1 - q_L)t + q_Lc = \beta$$

To see the result of proposition 4, notice that the settlement offers in Γ^a , $z(a = 1)$, $z(a = 0)$, are always higher or equal than the offers in Γ^{-a} . ■

Settlement in the admissible case

a1) Γ^a has an equilibrium which prescribes types sl and sh to apologize and P_F to forgive; P_L to offer $z(a = 1) = \lambda H + (1 - \lambda)L$, that sh accepts and sl rejects; P_F and P_L to offer $z(a = 0) = \lambda H$, that nh accepts and i, nl reject; J to award $m(a = 0) = 0$ and $m(a = 1) = L$ when uninformed, if and only if:

- i) $(1 - q_L)v_s - \beta > (q_L - \lambda)L$*
- ii) $\beta > (1 - q_L)\lambda H - q_L(1 - \lambda)L$*
- iii) $p_i > p_{nl}$*

a2) Γ^a has an equilibrium which prescribes types sl and sh to apologize and P_F to forgive; P_L to offer $z(a = 1) = \lambda H + (1 - \lambda)L$, that sh accepts and sl rejects; P_F and P_L to offer $z(a = 0) = \lambda H + (1 - \lambda)L$, that nh accepts and i, nl rejects; J to award $m(a = 0) = L$ and $m(a = 1) = L$ when uninformed, if and only if:

- i) $(1 - q_L)v_s - \beta > -(1 - q_L)L$*
- ii) $\beta > (1 - q_L)(\lambda H + (1 - \lambda)L)$*
- iii) $p_i < p_{nl}$*

a3) Γ^a has an equilibrium which prescribes type sh to apologize and P_F to forgive; P_L to offer $z(a = 1) = H$, that sh accepts; P_F and P_L to offer $z(a = 0) = \lambda H$, that nh accepts and i, nl, sl reject; J to award $m(a = 0) = 0$ and $m(a = 1) = H$ when uninformed, if and only if:

$$i) -\lambda L + q_L(\lambda L + (1 - \lambda)H) > (1 - q_L)v_s - \beta > (q_L - \lambda)H$$

$$ii) \beta > -(q_L - \lambda)H$$

$$iii) p_i > p_{nl} + p_{sl}$$

a4) Γ^a has an equilibrium which prescribes type sh to apologize and P_F to forgive; P_L to offer $z(a = 1) = H$, that sh accepts; P_F and P_L to offer $z(a = 0) = \lambda H + (1 - \lambda)L$, that nh accepts and i, nl, sl reject; J to award $m(a = 0) = L$ and $m(a = 1) = H$ when uninformed, if and only if:

$$i) -L + q_L(\lambda L + (1 - \lambda)H) > (1 - q_L)v_s - \beta > (q_L - \lambda)H - (1 - \lambda)L$$

$$ii) \beta > -(q_L - \lambda)H + (1 - \lambda)L$$

$$iii) p_i < p_{nl} + p_{sl}$$

Equilibria $a3, a4$ prescribes $m(a = 1) = H$. This holds for off-the-path beliefs of J when an apology is presented at trial specifying that it comes from type- h . Both equilibria also exist for off-the-path beliefs that imply $m(a = 1) = L$ if and only if $\lambda > q_L$.

Γ^a has no other equilibrium.

PROOF: Equilibrium $a1$. Let us solve it by backward induction. When J receives a case with apologies, she knows that it comes from type sl , the only type that has apologized and rejected the settlement offer, hence she awards $m(a = 1) = L$ when uninformed. On the other hand, when no apologies are presented, she awards $m(a = 0) = 0$ as types i and nl reject the offer, and $p_i > p_{nl}$.

Now let us consider the decision of D 's types that receive settlement offers. sl and sh apologize, and hence receive the offer $z(a = 1) = \lambda H + (1 - \lambda)L$ from L . The expected payoff of rejecting the offer and going to trial is $-L$ and $-\lambda H - (1 - \lambda)L$ for types sl and sh respectively; hence sl prefers to reject and sh is indifferent. On the other hand, types i, nl, nh do not apologize and receive the offer $z(a = 0) = \lambda H$. Then, by rejecting i, nh, nl will pay $0, \lambda L$, and λH respectively, making type nh willing to accept and types i and nl to reject the offer.

Then, using our LEMMA, the optimal settlement offers are $z(a = 0)$ and $z(a = 1)$, as these

leave indifferent the high–liability type in each case and low liability types reject it.

Now let us consider the decision to accept an apology and to apologize. As only types sh, sl apologize, F cannot deviate to reject it.

Finally, we need to consider the deviation payoffs of every guilty type. Types sl and sh payoff from apologizing are $(1 - q_L)v_s - \beta - q_L L$ and $(1 - q_L)v_s - \beta - q_L z(a = 1)$ respectively. Then notice that when considering deviation payoffs we need to compare the payoff of sh with the ones of nh , and the same holds for l types. To see this notice that the reasons for why types nh or nl accept and reject $z(a = 0)$ are only related to their respective liability levels. Hence, sl , after deviation, will reject $z(a = 0)$ and pay $m(a = 0)$ when J remains uninformed; $(1 - q_L)v_s - \beta > (q_L - \lambda)L$ implies that she cannot profitable deviate. On the other hand, sh will accept $z(a = 0) = \lambda H$ if she deviates. Then, if condition i is satisfied, she cannot profitable deviate as $-(z(a = 0) - q_L z(a = 1)) = -(1 - q_L)\lambda H + q_L(1 - \lambda)L < (q_L - \lambda)L < (1 - q_L)v_s - \beta$. In the same sense, $-\lambda L > -\beta - q_L L$ and $-\lambda H > -\beta - q_L(\lambda H + (1 - \lambda)L)$ imply that nl and nh cannot profitable deviate to apologize. Both inequalities are satisfied if condition ii holds.

Equilibrium a2. Let us solve it by backward induction. When J receives a case with apologies, she knows that it comes from type sl , the only type that has apologized and rejected the settlement offer, hence she awards $m(a = 1) = L$ when uninformed. On the other hand, when no apologies are presented, she awards $m(a = 0) = L$ as $p_i < p_{nl}$ and types i and nl reject the offer.

Now let us consider the decision of D 's types that receive settlement offers. sl and sh apologize, and hence receives the offer $z(a = 1) = \lambda H + (1 - \lambda)L$ from L . The expected payoff of rejecting the offer and going to trial is $-L$ and $-\lambda H - (1 - \lambda)L$ for types sl and sh respectively; hence sl prefers to reject and sh is indifferent. On the other hand, types i, nl, nh do not apologize and receive the offer $z(a = 0) = \lambda H + (1 - \lambda)L$. Then, by rejecting i, nl, nh will pay $(1 - \lambda)L, L$, and $\lambda H + (1 - \lambda)L$, making type nh willing to accept and types i and nl to reject the offer.

Then, using our LEMMA, the optimal settlement offers are $z(a = 0)$ and $z(a = 1)$, as these leaves indifferent the high–liability type in each case and low liability types reject it.

Now let us consider the decision to accept an apology and to apologize. As only types sl, sh apologize, F cannot deviate to reject it.

Finally, we need to consider the deviation payoffs of every guilty type. Types sl and sh payoff from apologizing are $(1 - q_L)v_s - \beta - q_LL$ and $(1 - q_L)v_s - \beta - q_Lz(a = 1)$ respectively. Then notice that when considering deviation payoffs we need to compare the payoff of sh with the ones of nh , and the same holds for l types. To see this notice that the reasons for why types nh or nl accept and reject $z(a = 0)$ are only related to their respective liability levels. Hence, sl , after deviation, will reject $z(a = 0)$ and pay $m(a = 0)$ when J remains uninformed; $(1 - q_L)v_s - \beta > -(1 - q_L)L$ implies that she cannot profitable deviate. On the other hand, sh will accept $z(a = 0) = \lambda H + (1 - \lambda)L$ if she deviates. Then, if condition i is satisfied, she cannot profitable deviate as $-(z(a = 0) - q_Lz(a = 1)) = -(1 - q_L)(\lambda H + (1 - \lambda)L) < -(1 - q_L)L < (1 - q_L)v_s - \beta$. In the same sense, $-L > -\beta - q_LL$ and $-\lambda H - (1 - \lambda)L > -\beta - q_L(\lambda H + (1 - \lambda)L)$ imply that nl and nh cannot profitable deviate to apologize. Both inequalities are satisfied if condition ii holds.

No other equilibrium. If the prior of n -types is higher than s 's ones then F will not forgive, so below we consider that the opposite holds. We need to prove that the following settlement is not possible:

- i) Settlement just with sl
- ii) Settlement with sl, nh
- iii) Settlement with sl, nl
- iv) Settlement with sh, nh
- v) Settlement with sh, sl, nh
- vi) Settlement with sh, sl, nl
- vii) Settlement with sh, sl, nh, nl

Let us prove i). If there is settlement with sl then it must be that $(1 - q_L)v_s - \beta - q_Lz(a = 1) > -\lambda L$ or $(1 - q_L)v_s - \beta - q_Lz(a = 1) > -L$ depending whether $p_i p_{nl}$. However, if sh does not apologize, it must be that $-\lambda H > (1 - q_L)v_s - \beta - q_Lz(a = 1)$ (or $-\lambda H - (1 - \lambda)L > (1 - q_L)v_s - \beta - q_Lz(a = 1)$). Then as $-\lambda L > -\lambda H$ (or $-L > -\lambda H - (1 - \lambda)L$) both conditions cannot be satisfied.

To see ii). Notice that if nh prefers to apologize it also must be true for sh .

On iii), it must be that $z(a = 1) = L$ as just low liability type settles. Then sh will pre-

fer to deviate, as $-\lambda H < (1 - q_L)v_s - \beta - q_L L$ given that if sl settles it must be that $-\lambda L < (1 - q_L)v_s - \beta - q_L L$.

Then assume iv). As only h -types apologize $z(a = 1) = H$. Then, as nh settles it must be $-\beta - q_L H > -\lambda L$ as types that do not apologize are l -types and innocent ones, using our LEMMA, $z(a = 0) = \lambda L$. Moreover, as nl does not apologize it must be that $-\lambda L > -\beta - q_L L$, where she reject $z(a = 1)$ and end at trial where J award $m = L$ for off-the-path beliefs that l -types end at trial when uninformed. Then as $-\beta - q_L H < -\beta - q_L L$, we cannot have this equilibrium. The same holds for off-the-path beliefs that specify $m(a = 1) = H$.

Now assume v. This is similar to iv. $z(a = 0) = \lambda L$, as only low liability and innocent ones do not apologize. $z(a = 1) = \lambda H + (1 - \lambda)L$. Then as nl does not apologize, $-\lambda L > -\beta - q_L L$, where we use that she will reject the offer $z(a = 0)$ if deviates and pays L at trial. On the other hand, if nh apologize, it must be that $-\beta - q_L z(a = 1) > -\lambda L$. However, as $z(a = 1) > L$ both inequalities cannot be satisfied.

Case vi. $z(a = 0) = \lambda H$, $z(a = 1) = \lambda H + (1 - \lambda)L$. As nh does not apologize it must be that $-\lambda H > \beta - q_L(\lambda H + (1 - \lambda)L)$. Then as nl apologize it must be that $-\beta - q_L L > -\lambda L$. Both inequalities cannot be satisfied.

Finally, vii. In this case, $z(a = 0) = 0$ and hence n types will prefer to deviate. ■

Settlement in the inadmissible case

a1) Γ^{-a} has an equilibrium which prescribes type s to apologize and P_F to forgive; P_L to offer $z(a = 1) = \lambda H$, that sh accepts and sl rejects; P_F and P_L to offer $z(a = 0) = \lambda H$, that nh accepts and i, nl reject; J to award $m = 0$ when uninformed, if and only if:

$$i) (1 - q_L)v_s - \beta > -(1 - q_L)\lambda L$$

$$ii) \beta > (1 - q_L)\lambda H$$

$$iii) p_i > p_{nl} + q_L p_{sl}$$

a2) Γ^{-a} has an equilibrium which prescribes type s to apologize and P_F to forgive; P_L to offer $z(a = 1) = \lambda H + (1 - \lambda)L$, that sh accepts and sl rejects; P_F and P_L to offer $z(a = 0) = \lambda H + (1 - \lambda)L$, that nh accepts and i, nl reject; J to award $m = L$ when unin-

formed, if and only if:

- i) $(1 - q_L)v_s - \beta > -(1 - q_L)L$
- ii) $\beta > (1 - q_L)(\lambda H + (1 - \lambda)L)$
- iii) $p_i < p_{nl} + q_L p_{sl}$

a3) Γ^{-a} has an equilibrium which prescribes type sh to apologize and P_F to forgive; P_L to offer $z(a = 1) = \lambda H$, that sh accepts; P_F and P_L to offer $z(a = 0) = \lambda H$, that nh accepts and i, nl, sl reject; J to award $m = 0$ when uninformed, if and only if:

- i) $-(1 - q_L)\lambda L > (1 - q_L)v_s - \beta > -(1 - q_L)\lambda H$
- ii) $\beta > (1 - q_L)\lambda H$
- iii) $p_i > p_{nl} + p_{sl}$

a4) Γ^{-a} has an equilibrium which prescribes type sh to apologize and P_F to forgive; P_L to offer $z(a = 1) = \lambda H + (1 - \lambda)L$, that sh accepts; P_F and P_L to offer $z(a = 0) = \lambda H + (1 - \lambda)L$, that nh accepts and i, nl, sl reject; J to award $m = L$ when uninformed, if and only if:

- i) $-(1 - q_L)L > (1 - q_L)v_s - \beta > -(1 - q_L)(\lambda H + (1 - \lambda)L)$
- ii) $\beta > (1 - q_L)(\lambda H + (1 - \lambda)L)$
- iii) $p_i < p_{nl} + p_{sl}$

PROOF: Equilibrium a1. Let us solve it by backward induction. When J receives a case she knows that it comes from type sl , the only type that has apologized and rejected the settlement offer $z(a = 1)$, tendered by L , or from type nl and type i that reject $z(a = 0)$ offered by F and L ; hence she awards $m = 0$ when uninformed given $p_i > p_{nl} + q_L p_{sl}$. Now let us consider the decision of D 's types that receive settlement offers. sl and sh apologize, and hence receives the offer $z(a = 1) = \lambda H$ from L . The expected payoff of rejecting the offer and going to trial is $-\lambda L$ and $-\lambda H$ for types sl and sh respectively; hence sl prefers to reject and sh is indifferent. On the other hand, types i, nl, nh do not apologize and receive the offer $z(a = 0) = \lambda H$. Then, by rejecting i, nl, nh will pay 0, λL , and λH ,

making type nh willing to accept and types i and nl to reject the offer.

Then, using our LEMMA, the optimal settlement offers are $z(a = 0)$ and $z(a = 1)$, as these leave indifferent the high-liability type in each case and low liability types reject it.

Now let us consider the decision to accept an apology and to apologize. As only types sh, sl apologize, F cannot deviate to reject it.

Finally, we need to consider the deviation payoffs of every guilty type. Types sl and sh payoff from apologizing are $(1 - q_L)v_s - \beta - q_L\lambda L$ and $(1 - q_L)v_s - \beta - q_Lz(a = 1)$ respectively. Then notice that when considering deviation payoffs we need to compare the payoff of sh with the ones of nh , and the same holds for l types. To see this notice that the reasons for why types nh or nl accept and reject $z(a = 0)$ are only related to their respective liability levels. Hence, sl , after deviation, will reject $z(a = 0)$ and pay $m = 0$ when J remains uninformed; $(1 - q_L)v_s - \beta > -(1 - q_L)\lambda L$ implies that she cannot profitable deviate. On the other hand, sh will accept $z(a = 0) = \lambda H$ if she deviates. Then, if condition i is satisfied, she cannot profitable deviate as $-(z(a = 0) - q_Lz(a = 1)) = -(1 - q_L)\lambda H < -(1 - q_L)\lambda L < (1 - q_L)v_s - \beta$. In the same sense, $-\lambda L > -\beta - q_L\lambda L$ and $-\lambda H > -\beta - q_L\lambda H$ imply that nl and nh cannot profitable deviate to apologize. Both inequalities are satisfied if condition ii holds.

Equilibrium a2. Let us solve it by backward induction. When J receives a case she knows that it comes from type sl , the only type that has apologized and rejected the settlement offer $z(a = 1)$, tendered by L , or from type nl and type i that rejects $z(a = 0)$ offered by F and L ; hence she awards $m = L$ when uninformed given $p_i < p_{nl} + q_L p_{sl}$.

Now let us consider the decision of D 's types that receive settlement offers. sl and sh apologize, and hence receives the offer $z(a = 1) = \lambda H + (1 - \lambda)L$ from L . The expected payoff of rejecting the offer and going to trial is L and $\lambda H + (1 - \lambda)L$ for types sl and sh respectively; hence sl prefers to reject and sh is indifferent. On the other hand, types i, nl, nh do not apologize and receive the offer $z(a = 0) = \lambda H + (1 - \lambda)L$. Then, by rejecting i, nl, nh will pay $(1 - \lambda)L$, L , and $\lambda H + (1 - \lambda)L$, making type nh willing to accept and types i and nl to reject the offer.

Then, using our LEMMA, the optimal settlement offers are $z(a = 0)$ and $z(a = 1)$, as these leave indifferent the high-liability type in each case and low liability types rejects it.

Now let us consider the decision to accept an apology and to apologize. As only types ss

apologize, F cannot deviate to reject it.

Finally, we need to consider the deviation payoffs of every guilty type. Types sl and sh payoff from apologizing are $(1 - q_L)v_s - \beta - q_L\lambda L$ and $(1 - q_L)v_s - \beta - q_Lz(a = 1)$ respectively. Then notice that for deviation payoffs we need to compare the payoff of sh with the ones of nh , and the same hold for l types. To see this notice that the reasons for why types nh and nl accepts and rejects $z(a = 0)$ are only related to their respective liability levels. Hence, sl , after deviation, will reject $z(a = 0)$ and pay $m = L$ at trial; $(1 - q_L)v_s - \beta > -(1 - q_L)L$ implies that she cannot profitable deviate. On the other hand, sh will accept $z(a = 0)$ if she deviates. Then, if condition i is satisfied, she cannot profitable deviate as $-(z(a = 0) - q_Lz(a = 1)) = -(1 - q_L)(\lambda H + (1 - \lambda)L) < -(1 - q_L)L < L$. In the same sense, $-L > -\beta - q_LL$ and $-(\lambda H + (1 - \lambda)L) > -\beta - q_L(\lambda H + (1 - \lambda)L)$ imply that nl and nh cannot profitable deviate to apologize. Both inequalities are satisfied if condition ii holds.

No other equilibrium. If the prior of n -types is higher than s 's ones then F will not forgive, so below we consider that the opposite holds. We need to prove that the following settlement is not possible:

- i) Settlement just with sl
- ii) Settlement with sl, nh
- iii) Settlement with sl, nl
- iv) Settlement with sh, nh
- v) Settlement with sh, sl, nh
- vi) Settlement with sh, sl, nl
- vii) Settlement with sh, sl, nh, nl

Let us prove i). If there is settlement with sl then it must be that $(1 - q_L)v_s - \beta - q_Lz(a = 1) > -\lambda L$ or $(1 - q_L)v_s - \beta - q_Lz(a = 1) > -L$ depending whether $p_i p_{nl}$. However, if sh does not apologize, it must be that $-\lambda H > (1 - q_L)v_s - \beta - q_Lz(a = 1)$ (or $-\lambda H - (1 - \lambda)L > (1 - q_L)v_s - \beta - q_Lz(a = 1)$). Then as $-\lambda L > -\lambda H$ (or $-L > -\lambda H - (1 - \lambda)L$) both conditions cannot be satisfied.

To see ii). Notice that if nh prefers to apologize it also must be true for sh .

On iii, it must be that $z(a = 1) = \lambda L$: just l -types apologize and when uninformed J awards $m = 0$ as h -types accept $z(a = 0) = \lambda H$. Then sh will prefer to deviate as $-\lambda H < (1 - q_L)v_s - \beta - q_L\lambda L$ given that if sl settles it must be that $-\lambda L < (1 - q_L)v_s - \beta - q_L\lambda L$. Then assume iv). As only h -types apologize, $z(a = 1) = \lambda H$. Then, as nh settles it must be $-\beta - q_L\lambda H > -\lambda L$ as types that do not apologize are l -types and innocent ones, thus using our LEMMA, $z(a = 0) = \lambda L$. Moreover, as nl does not apologize it must be that $-\lambda L > -\beta - q_L\lambda L$; where we use the fact that she will reject $z(a = 1)$ and end at trial where J award $m = 0$, as the equilibrium prescribes that just i rejects an offer. Then, as $-\beta - q_L\lambda H < -\beta - q_L\lambda L$, we cannot have this equilibrium.

Now assume v. This is similar to iv. Assume $p_i > p_{sl}$. $z(a = 0) = \lambda L$, as only l -types and innocent ones do not apologize. $z(a = 1) = \lambda H$. Then as nl does not apologize, $-\lambda L > -\beta - q_L\lambda L$, where we use that she will reject the offer $z(a = 0)$ if deviates and pays λL at trial. On the other hand, if nh apologizes, it must be that $-\beta - q_Lz(a = 1) > -\lambda L$. However, as $z(a = 1) > \lambda L$ both inequalities cannot be satisfied. The same argument holds for $p_i < p_{sl}$

Case vi. Assume $p_i > p_{sl} + p_{nl}$. $z(a = 0) = \lambda H = z(a = 1)$. As nh does not apologize it must be that $-\lambda H > -\beta - q_L(\lambda H)$. Also, as nl apologize it must be that $-\beta - q_L\lambda L > -\lambda L$. Both inequalities cannot be satisfied. Similar for $p_i < p_{sl} + p_{nl}$ and $z(a = 0) = \lambda H + (1 - \lambda)L = z(a = 1)$. Finally vii. F will drop the case after receiving no apologies, as only i refrains to tender it. Given this, type nh will prefer to deviate to not apologies as $-\beta - q_Lz(a = 1) < -q_Lz(a = 0)$ given that $z(a = 1) = \lambda H + (1 - \lambda)L > (1 - \lambda)L = z(a = 0)$. ■

References

- Johannes Abeler, Juljana Calaki, Kai Andree, and Christoph Basek. The power of apology. *Economics Letters*, 107:233–235, 2010.
- Lucian Bebchuk. Litigation and settlement under imperfect information. *RAND Journal of Economics*, 15(3):404–415, 1984.
- Shereen J. Chaudhry and George Loewenstein. Thanking, apologizing, bragging, and blaming: Responsibility exchange theory and the currency of communication. *Psychological Review*, 126(3):313–344, 2019.
- Jonathan Cohen. Legislating apology: The pros and cons. *University of Cincinnati Law Review*, 70:819–872, 2002.
- A. Daughety and J. Reinganum. Settlement. *The Encyclopedia of Law and Economics (Second Ed.) – Ch 15, Vol. 8 (Procedural Law and Economics, Ed. by Chris William Sanchirico)*, published by Edward Elgar Publishing Ltd., pages 386–471, 2012.
- Andrew Daughety and Jennifer Reinganum. Keeping society in the dark: On the admissibility of pretrial negotiations as evidence in court. *RAND Journal of Economics*, 26(2): 203–221, 1995.
- Urs Fischbacher and Verena Utikal. On the acceptance of apologies. *Games and Economic Behavior*, 82:592–608, 2013.
- Thomas Gallagher, Amy D. Waterman, Alison C. Ebers, Victoria J. Fraser, and Wendy Levinson. Physicians are reluctant to give apologies because it damages their reputation. *JAMA, the Journal of the American Medical Association*, 289(8):1001–1007, 2003.
- Dwight Golann. Dropped medical malpractice claims: Their surprising frequency, apparent causes, and potential remedies. *Health Affairs*, 30(7):1343–1350, 2011.
- Basil Halperin, Benjamin Ho, John A. List, and Ian Muir. Towards an understanding of the economics of apologies: evidence from a large-scale natural field experiment. 2019.

- Benjamin Ho. Apologies as signals: With evidence from a trust game. *Management Science*, 58(1):141–158, 2012.
- Benjamin Ho and Elaine Liu. Does sorry work? the impact of apology laws on medical malpractice. *Journal of Risk and Uncertainty*, 43(4):141–167, 2011.
- Peter H. Kim, Kurt T. Dirks, Cecily D. Cooper, and Donald L. Ferrin. When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation. *Organizational Behavior and Human Decision Processes*, 99(1):49–65, 2006.
- John Kleefeld. Promoting and protecting apologetic discourse through law: A global survey and critique of apology legislation and case law. *Oñati Socio-legal Series*, 7(3): 455–496, 2017.
- Benjamin J. McMichael. Insuring apologies. *Florida State University Law Review*, 48, 2021.
- Benjamin J. McMichael, R. Lawrence Van Horn, and W. Kip Viscusi. “sorry” is never enough: How state apology laws fail to reduce medical malpractice liability risk. *Stanford Law Review*, 71(2):341–409, 2019.
- Cayce Myers. Review apology, sympathy, and empathy: The legal ramifications of admitting fault in u.s. public relations practice. *Public Relations Review*, 42:176–183, 2016.
- Ken-Ichi Ohbuchi and Kobun Sato. Children’s reactions to mitigating accounts: Apologies, excuses, and intentionality of harm ken-ichi ohbuchi & kobun sato. *The Journal of Social Psychology*, 134(1):5–17, 1994.
- Yohsuke Ohtsubo and Esuka Watanabe. Do sincere apologies need to be costly? test of a costly signaling model of apology. *Evolution and Human Behavior*, 30(2):114–123, 2009.
- Daniel J. Seidmann. The effects of a right to silence. *The Review of Economic Studies*, 72 (2):593–614, 2005.
- Kathryn Spier. Litigation. *Handbook of law and economics*, 1:259–342, 2007.

Prue Vines. Apologising for personal injury in law: Failing to take account of lessons from psychology in blameworthiness and propensity to sue. *Psychiatry, Psychology and Law*, 22(4):624–634, 2015.

Bernard Weiner, Sandra Graham, Orli Peter, and Mary Zmuidinas. Public confession and forgiveness. *Journal of Personality*, 59(2):281–312, 1991.