Dispute Settlement with Second-Order Uncertainty

Mostafa Beshkar
Indiana University

Jee-Hyeong Park*
Seoul National University

June 2019

Abstract

The literature on pre-trial dispute settlement has focused on the effect of first-order uncertainty on pre-trial settlement bargaining while assuming common knowledge about higher-order beliefs. We study the effect of uncertainty regarding higher-order beliefs and show that ignorance about higher-order beliefs improves the efficiency of settlement bargaining by raising the disputing parties' expected payoffs in bargaining. We introduce uncertainty about higher-order beliefs by assuming that one player receives a private and imperfect signal of another player's private type. We show that such signals could improve the efficiency of settlement bargaining only if they are privately observed: the informational value associated with the signal completely disappears if it is publicly observable.

*Beshkar: Associate Professor of Economics, Indiana University (mbeshkar@indiana.edu). Park: Professor of Economics, Seoul National University (j-hpark@snu.ac.kr). For their useful comments, we would like to thank Treb Allen, Pol Antras, Kyle Bagwell, Andrew Bernard, Eric Bond, Andrew Daughety, Taiji Furusawa, Jota Ishikawa, Gita Gopinath, Jihong Lee, Marc Melitz, Hiroshi Mukunoki, Jennifer Reinganum, Joel Rodrigue, Kenneth Rogoff, Kathryn Spier, Christopher Snyder, Robert Staiger, Alexander Wolitzky and the seminar participants at Dartmouth College, Harvard, Stanford, Vanderbilt, the Asia Pacific Trade Seminar, and the Midwest Theory Conference. For their detailed and constructive comments, we are grateful to the editor and anonymous referees. We also thank Jun-Tae Park and Sang Min Lee for their research assistance. Park would like to thank the people at the economics department of Harvard for their hospitality during his visit (2016-2017) as a Fulbright scholar, conducting this research. Park acknowledges the financial support from National Research Foundation of Korea (NRF-2014-S1A3A2043505), Seoul National University Asia Center (Asia Research Foundation Grant #SNUAC-2017-008), and 'Overseas Training Expenses for Humanities & Social Sciences' through Seoul National University.

1 Introduction

Pre-trial dispute settlement has been often studied in the economics literature as a bargaining game under asymmetric information. The information asymmetry considered in the literature is limited to the uncertainty about the first-order beliefs, while higher-order beliefs are assumed to be common knowledge. In particular, it is commonly assumed that each disputing party possesses private information about its type (determining its belief of court outcomes) while the distribution of types is commonly known to both parties.[1]

Our objective in this paper is to propose a model of pre-trial settlement bargaining under higher-order uncertainty. To this end, we analyze a settlement bargaining game in which one player receives a *private and imperfect signal* of another player's private type, thereby generating second-order uncertainty. In particular, we study a game of bilateral bargaining over actions in which *(i)* the outcome of litigation (or arbitration) depends on the informed party's *type* and *(ii)* the uninformed party receives an *imperfect private signal* about the other party's type.[2]

We first analyze a simple pre-trial bargaining game of the following structure. At the beginning of the game a defendant observes her type (either high or low) and a complainant receives an imperfect signal about the defendant's type. The defendant then proposes an action to take for settlement, which may be accepted or rejected by the complainant. In the case of rejection, the case is litigated, with the outcome of court ruling being uncertain: A high-type defendant would expect a more favorable ruling than a low-type one but the court outcome is still uncertain as the court ruling is based on its own imperfect signal of the defendant's type.[3]

The high-type defendant benefits more from taking a higher level action than the low type, but a higher level action generates a larger negative effect on the

[1]There is a growing literature, including Bergemann and Morris (2009), Chen et al. (2017), Morris et al. (2016), that analyzes the effect of higher order beliefs and associated uncertainty on games and mechanism design problems. The literature on pre-arbitration settlement, however, has not explored such an issue. For a comprehensive review of the literature on litigation and pre-arbitration settlement see Daughety and Reinganum (2017) and Spier (2007).

[2]Our study is distinct from the literature on two-sided private information. Schweizer (1989) assumes that each disputing party receives an independent signal about the probability of its success in the court. Daughety and Reinganum (1994) also propose a dispute settlement model with two-sided imperfect information under which the complainant is privately informed about the extent of damages incurred, and the defendant is privately informed about the likelihood of being found liable for damages in the court. In both of these studies, the distribution of types is common knowledge and thus no second-order uncertainty exists.

[3]This bargaining game is a signaling game in which the informed party tries to signal her

complainant's payoff. While there exists a set of joint-payoff maximizing actions that assigns a higher level action for high-type realization, attaining the first-best outcome may not be incentive compatible as the low-type defendant has an incentive to mimic the high type. To deter such an opportunistic behavior, the complainant needs to reject the proposal of a presumably high-type defendant with a positive probability, resorting to an action determined by a court (or an arbitrator) despite the existence of a mutually beneficial settlement action.[4]

Our main finding is related to the role of uncertainty about second-order beliefs on the likelihood and efficiency of settlement. Uncertainty about second-order beliefs depends on whether the imperfect signal that the complainant receives is publicly observable or not. If the signal is public—implying no uncertainty in second-order beliefs—then the signal has no bearing on the equilibrium of the game. The imperfect signal becomes useful (by way of reducing the likelihood of litigation) only if it is privately observed by the complainant. In other words, we find that *ignorance* about higher-order beliefs improves the efficiency of settlement bargaining by raising the disputing parties' expected payoffs in bargaining.[5]

For a general intuition for this result, note the difference between litigation strategy of the complainant under private and public signals. If the signal is public, the defendant ends up confronting the same risk of litigation regardless of its type. While a revealed public signal changes the common prior of the parties, it does not affect the complainant's rejection probability that is required to deter the low type's opportunistic behavior of mimicking the high type's settlement pro-

type through her take-it-or-leave-it settlement offer. We also model the settlement negotiation as a screening game under which the uninformed party (i.e., the complainant in our model) makes a settlement offer. We find that in a screening game, the second-order uncertainty does not play any role. The best offer strategy of the complainant depends only on the probability that the defendant is a high type conditional on its signal, which is unaffected by whether its signal is private or public.

[4] Given that the complainant knows the defendant's realized type (which occurs in a fully-separating equilibrium), there exists an range of settlement actions that are preferred by both parties to the uncertain outcome of arbitration in the pre-trial bargaining game that we analyze.

[5] The result that common knowledge can reduce the efficiency of bargaining or hinder negotiation arises in other settings as well. Ayres and Nalebuff (1996) provide a series of examples showing how mediators can facilitate agreement by preventing the creation of common knowledge. They argue that preserving ignorance about higher-order information—which may be achieved by employing a mediator who transmits only first-order information— can promote trade between a buyer and a seller. Another related result is the *anti-public-signal* result of Morris and Shin (2002), which will be discussed below.

posal.[6] In contrast, if the defendant does not observe the signal received by the complainant, i.e., in the presence of second-order uncertainty, then it is possible to have the opportunistic low-type defendant face a higher likelihood of litigation than the high-type defendant. This is precisely why a private signal could change (and improve) the equilibrium while a public signal has no effect on the equilibrium.

The Dispute Settlement Process (DSP) of the World Trade Organization (WTO) provides an example of bargaining under asymmetric information with higher-order uncertainty. The obligations of an importing country under the WTO are *contingent* on its domestic political economy conditions, which are likely to be the private information of the importing government. Other governments, however, could also conduct their own investigations and receive informative signals about the importing country's political economy conditions. These signals, which are potentially the private information of the investigating governments, creates second-order uncertainty in the pre-arbitration dispute settlement game.[7]

To analyze the pre-arbitration settlement negotiation of a WTO trade dispute, we extend our simple model with one action variable (of the defendant) into a model with two action variables, allowing also the complainant's action to be a subject of settlement as well as the defendant's. Governments in a trade dispute often negotiate over a bilateral change in their protection levels, as discussed in Section 5.2 on a specific WTO dispute case.[8] Being similar to the defendant's action, the complainant may also benefit from taking a higher level action (higher protection), but a higher level action of the complainant generates a larger negative effect on the defendant's payoff. The defendant then can offer a combination of actions to take as settlement, proposing a tolerable level of the complainant's withdrawal of concession previously granted to the defendant as a price for taking

[6]Due to the typical multiplicity of equilibria that arises in the signaling game, we invokes a refinement criterion often employed in the pre-trial settlement bargaining game, namely *Divinity* refinement. This enables us to focus on the separating equilibrium that maximizes the high-type defendant's expected payoff.

[7]For example, exporting firms may have some cost information that are common among firms in the same industry, which in turn can be informative in accessing the level of damages inflicted on import-competing firms in their export destination. Such cost related information is often confidential business information, as illustrated by a large number of "Best Information Available" cases in the U.S. anti-dumping investigations caused by refusing to submit cost-related information despite the risk of paying excessively high anti-dumping duties. In the absence of exporting companies' submission of cost-related information, the US Department of Commerce calculates its dumping margins based on "Best Information Available," often the estimated costs of such exporters conjectured by the import-competing firms who filed an anti-dumping petition.

[8]We follow the literature on trade agreements by assuming that intergovernmental transfers

a higher level action (protection) of her own.[9]

In contrast to the simple bargaining model with one action variable, this model with two action variables may entail Pareto-inefficient actions (more precisely, action combinations) as well as Pareto-efficient ones as a settlement outcome.[10] Our analysis shows that under high protectionist pressures at home, a government will propose tariff levels that are higher than the Pareto-efficient protection levels. Choosing such an inefficient action combination is more costly for a low type than for a high one, which enables the high type to signal her type through such an offer, generating a fully separating equilibrium.[11] The proposed tariff levels are possibly even higher than the ones that Dispute Settlement Body (arbitrator) of the WTO would recommend when it finds evidence in favor of the defending government. This finding provides a new perspective on the observation that the DSB often rules against a defending government, recommending reduction or removal of the contingent protection.[12]

Despite of having a Pareto-inefficient action combination as a settlement outcome, allowing governments to negotiate not only over the defendant's protection measures but also over the complainant's withdrawal of concessions improves the ex ante joint payoffs of disputing parties.[13] Such a gain largely comes from reducing the complainant's settlement rejection probability required to deter the low-type defendant's mimicking the high type's proposal, which in turn decreases the likelihood of invoking a costly litigation: The high-type defendant can more effectively signal her type with two action variables than with one action vari-

are usually in the form of policy adjustments, such as a bilateral change in the level of protection, rather than cash transfers. This assumption reflects some realities about international relations. First, transferring cash among governments involves both political and public finance costs, which reduces its appeal as a compensation mechanism. Moreover, in response to violation of their rights in an international agreement, governments normally seek compensations through withdrawal of concessions previously granted to the defecting government. This *self-help* method of receiving compensation may reflect the fact that the defecting country, being a sovereign state, cannot be coerced to compensate the injured government.

[9] Bagwell and Staiger (2005) show that when cash transfer is possible, the governments could implement the first-best outcome by requiring a proper amount of cash transfers as a price for imposing a contingent tariff.

[10] In our one-action-variable model, all relevant actions are Pareto-efficient as a higher level action benefits the defendant but hurts the complainant.

[11] As the receiver's private signal of the sender's type becomes increasingly accurate, the high type's equilibrium offer will approach a Pareto efficient action combination.

[12] Sykes (2003) points out that the arbitrator has always ruled against the defending party in litigation regarding safeguard measures.

[13] The complainant's ex ante payoff is unaffected by such a change in action variables, as the defendant takes all the informational rent through her take-it-or-leave-it settlement offer.

able in her settlement proposal. The WTO dispute settlement protocol already reflects this desirability of allowing governments to negotiate over bilateral change in their protection levels. While the WTO focus on determining whether the defendant's protection measure violates its rules or not in its arbitration process, it does not provide any specific guideline for disputing parties in the settlement stage, except emphasizing that it prefers settlement over litigation.

In a similar manner, the WTO indirectly endorses having the complainant information remain private (instead of requiring such information be publicized) in the settlement stage, a preferred institutional arrangement according to our *anti-public-signal* result on the settlement of trade disputes. Such a non-transparent settlement stage is in contrast with the WTO's arbitration process that is highly public and transparent. On the enforcement of international trade agreements, especially at the stage of determining whether to invoke retaliation against potential violations, Park (2011) provides a *pro-public-monitoring* result. Using a repeated-game framework with imperfect monitoring of the potential use of concealed trade barriers, Park (2011) demonstrates that publicizing the imperfect private signal of potential deviations may facilitate a higher level of cooperation by relaxing the incentive constraint associated with utilizing imperfect private signals in invoking punishment.[14] Thus, our analysis together with Park (2011) provide an explanation for why the WTO may take very different stances in the pre-arbitration stage and in the arbitration stage with regard to the publicity of information utilized in such procedures.

Our analysis also predicts that an improvement in the quality of signals received by the complaining government will reduce the probability of litigation. This theoretical finding is consistent with the evidence provided by Ahn, Lee, and Park (2014) who find a positive correlation between a proxy for information asymmetry and rate of litigation. The fact that the rate of the WTO disputes have decreased over time may also reflect a reduction in information asymmetry between the parties (i.e., improved signals) after years of partnership.[15]

[14]Using a repeated game framework with incomplete information of potentially persistent political pressure for protection, Bagwell (2009) analyzes enforcement issues in trade agreements, demonstrating that a government facing a low political pressure may "pool" and apply its tariff at the bound rate, which is inefficiently high for her.

[15]The number of WTO dispute cases decreased from 335 during its first 10 years (1995-2005) to 165 during the next 10 years (2006-2015). This decrease in the WTO disputes is even more surprising once we consider the steady expansion of the WTO membership from to 123 countries in 1995 to 162 countries in 2015, including major ones, such as China (2001) and Russia (2012). As trading partners interact for a longer period, informational asymmetry between them will naturally decrease.

Our paper is closely related to the recent literature on dispute settlement in the WTO. In particular, Beshkar (2010b, 2016), Park (2011), and Maggi and Staiger (2017, 2018) study the role of the WTO as a *public signaling* device that reveals some useful —albeit imperfect— information about the type or action of the defending party.[16] A question has been hovering over these studies: How would the value of the court as a public signaling device change if we consider the ability of the uninformed parties to conduct their own investigations to obtain an independent signal about the state of the world or the other party's actions? This question is particularly interesting given that the disputing parties are most likely better equipped than the WTO arbitrators to monitor and extract information about the private type or actions of each other. Our study advances this line of research by shedding light on the impact of private monitoring conducted by the disputing parties.

Outside the literature on disputes and settlement, our anti-public-signal result may be compared to that of Morris and Shin (2002) who show that an increase in the precision of public information may generate a detrimental effect on the overall welfare of participants in a coordination game when each participant has access to private information. The public information in Morris and Shin (2002) serves as a coordination device among participants, creating the possibility of inducing a weight on the public information that is higher than the socially optimal level. In our signaling game of settlement bargaining, the public information eliminates second-order uncertainty that enables the receiver to make its rejection threat contingent upon the information about fundamentals (i.e., the sender's type), which in turn completely eliminates its informational value.

Our result that a public signal has no impact on the equilibrium is related to the analysis of Bagwell (1995). He shows that any level of noise in a follower's observation of a first mover's action can induce the follower to completely ignore its imperfect information in a pure strategy equilibrium, which in turn eliminates the first mover advantage.[17] Public information without any noise will induce the players to utilize such information in our settlement bargaining game, eliminating the need for inefficient litigation in the equilibrium. In contrast to the game analyzed by Bagwell (1995), the imperfect information of the follower (settlement offer recipient) retains its informational value as long as it generates second-order

[16] Another related paper is Maggi and Staiger (2011) in which the arbitrator is modeled as an arbitrator that interprets ambiguous obligations, fills gaps in the agreement, and modifies rigid obligations. See Park (2016) for a comprehensive review of the recent literature on trade disputes and settlement.

[17] Maggi (1999) demonstrates that the strategic value of commitment (e.g., moving first) is re-

uncertainty to the first mover (settlement offer maker).

Applying mechanism design to international conflict resolution, Hörner, Morelli, and Squintani (2015) demonstrate how a mediator without enforcement power can replicate the welfare outcome of an optimal settlement mechanism that utilizes an arbitrator with enforcement power. Under their model, the mediator overcomes its lack of enforcement power by choosing a recommendation strategy that does not reveal the type of a weak player to a strong player. This restrains the strong player's incentive for fighting against the weak one. One could interpret the optimality of uncertainty in the mediator's recommendation as optimality of second order uncertainty as we find in this paper.

In terms of informational structure, Feinberg and Skrzypacz (2005) analyze a similar bargaining game in which a seller has private information about his beliefs about the buyer's private valuation (i.e., type), thus entailing second-order uncertainty. This second-order uncertainty creates a surprising result that delay in bargaining occurs even when a rational seller makes frequent offers to a rational buyer with common knowledge of gains from trade, thus generating a result that does not follow the "Coase property."

In Section 2, we describe the basic setup of our pre-trial bargaining model with one action variable. Section 3 analyzes the settlement bargaining as a signaling game with imperfect private signals. In Section 4, we shed light on the role of second-order uncertainty in settlement bargaining by analyzing the effect of publicizing the complainant's signal as well as the effect of modeling dispute settlement as a screening, rather than signaling model. Section 5 extends our analysis to a pre-trial bargaining model with two action variables to analyze the WTO dispute settlement. Finally, in Section 6, we provide some concluding remarks.

2 Basic Setup

There are two parties (D)efendent and (C)omplainant, with D having an action variable $\tau \in R^{+}$. The parties' payoffs are denoted by $W^{D}(\tau;\theta)$ and $W^{C}(\tau)$, respectively, where, θ represents the state of the world that affects D's payoff (and only D's). We assume that:

stored even with imperfect observability of commitment when a leader has private type information. As the optimal leader action depends on her type, the follower has an incentive to utilize even its imperfect information of the leader's action, restoring at least a certain value of commitment. Although our settlement bargaining game also analyzes the situation in which the first mover (a defendant who makes a take-it-or-leave-it offer) has private type information, the follower's imperfect information is not about the first mover's action but about her type.

1. D's payoff is concave and initially increasing in its own action:

$$\frac{\partial^2 W^D}{\partial \tau^2} < 0,$$
$$\left.\frac{\partial W^D}{\partial \tau}\right|_{\tau=0} > 0.$$

2. D's action has a negative externality on the other party:

$$\frac{\partial W^C}{\partial \tau} < 0.$$

3. The marginal payoff of D from her own action is increasing in the state parameter, θ :

$$\frac{\partial W^D}{\partial \tau d\theta} > 0.$$

4. The sum of payoffs is concave in τ:

$$\frac{\partial^2 \left(W^D + W^C\right)}{\partial \tau^2} < 0$$

Letting $\tau^N \equiv \arg\max W^D$ and $\tau^E \equiv \arg\max_\tau \left(W^D + W^C\right)$ denote, respectively, the non-cooperative and jointly-efficient levels of τ, these assumptions imply that

$$\tau^N(\theta) > \tau^E(\theta)$$

We further assume that the state of the world can take one of two levels: high (h) and low (l), where $h > l$. Therefore,

$$\tau^N(h) > \tau^N(l),$$
$$\tau^E(h) > \tau^E(l).$$

Finally, we assume that

$$\tau^N(l) > \tau^E(h).$$

This final assumption simplifies the analysis by eliminating the possibility of *commitment overhang* under an optimal agreement.[18] The above assumptions on payoff functions represent the situation in which the state contingent jointly-efficient

[18]For an analysis of commitment overhang, see Amador and Bagwell (2013); Beshkar et al. (2015); Beshkar and Bond (2017).

action is different from (lower than) the non-cooperative action. This leads to inefficient actions in the absence of complete information of the state of world, as shown below.

Information Structure While D observes θ privately, C receives an imperfect *private* signal of θ, denoted by θ^C, which is accurate with a probability of γ, namely,

$$\Pr\left(\theta^C = l\,|\theta = l\right) = \Pr\left(\theta^C = h\,|\theta = h\right) = \gamma \in \left(\frac{1}{2}, 1\right).$$

The arbitrator, A, also receives a signal, denoted by θ^A, with accuracy γ^A, namely,

$$\Pr\left(\theta^A = h|\theta = h\right) = \Pr\left(\theta^A = l|\theta = l\right) = \gamma^A \in \left(\frac{1}{2}, 1\right).$$

We do not make any assumption regarding the relative accuracy of the signals observed by the complainant and the arbitrator, i.e., γ and γ^A. We, however, assume that these signals are independent.

Arbitration

We model the arbitrator as a court that, if invoked, issues a binding ruling as a function of its informative signal, θ^A. The arbitrator's ruling is an action $\tau^A\left(\theta^A\right)$ to be implemented by D. The objective of the court is to maximize the expected joint payoff of the parties given its observed signal, θ^A :

$$\tau^A\left(\theta^A\right) \equiv \arg\max_{\tau} \sum_{\theta=l,h} \Pr\left(\theta|\theta^A\right)\left[W^D\left(\tau;\theta\right) + W^C\left(\tau\right)\right].$$

In calculating the conditional probability of D's type, the arbitrator utilizes only its own observed signal, θ^A, and the common prior of D being a high type, denoted by $\rho \in (0,1)$. Thus, the arbitrator ignores (i.e., is ignorant of) what happened in the settlement bargaining process. Then, we can obtain the following characterization of arbitrated actions.

Lemma 1. $\tau^E\left(l\right) < \tau^A\left(l\right) < \tau^A\left(h\right) < \tau^E\left(h\right)$

Note that optimal τ^A must be between $\tau^E\left(l\right)$ and $\tau^E\left(h\right)$. Moreover, optimal τ^A will be closer to $\tau^E\left(h\right)$, the higher is $\Pr\left(\theta = h|\theta^A\right)$. Therefore, since $\Pr\left(\theta = h|\theta^A = h\right) > \Pr\left(\theta = h|\theta^A = l\right)$, we must have $\tau^A\left(h\right) > \tau^A\left(l\right)$.

Letting $W^i_A(\theta)$ denote the expected payoff of a party $i = \{D, C\}$ from arbitration when the true state of the world is θ, and D_θ denote a type-θ defendant, it is straightforward to show that

Corollary 1. *The expected joint payoff under optimal arbitration is strictly higher than the expected joint payoff under any settlement without arbitration. Specifically, we have*

$$E_\theta\left[W^D_A(\theta) + W^C_A(\theta)\right] > E_\theta\left[W^D(\tau;\theta) + W^C(\tau)\right] \tag{1}$$

for all $\tau \in R^+$.

If there exists τ that violates the above inequality, then such τ should be preferred over the arbitration assigning $\tau^A(l)$ and $\tau^A(h)$ depending on θ^A, thus invalidating Lemma 1.

With regard to potential benefit from settlement over arbitration, we can obtain the following lemma:

Lemma 2. *For any realization of θ, there are mutual gains from settlement in lieu of arbitration, that is,*

$$\forall\theta \in \{l, h\}, \exists\tau : W^D(\tau,\theta) > W^D_A(\theta) \text{ and } W^C(\tau,\theta) > W^C_A(\theta). \tag{2}$$

Lemma 2 reflects both the fact that the outcome of arbitration is uncertain due to imperfectness in the arbitrator's information, and the fact that the concavity of D's payoff and the joint payoff functions in τ implies risk aversion of bargaining parties. Note that the optimality of the arbitration process is only a sufficient—but not necessary– condition for inequalities in (1) and (2) to hold. In fact, optimal arbitration is not necessary for our subsequent analysis either: any non-strategic (i.e. ignoring what happened in the settlement bargaining stage) arbitration system, or more generally any non-settlement contingency, of which the expected outcome yielding these inequality relations will work.[19]

As the first step in analyzing the pre-arbitration bargaining game, we determine each party's *outside* option based on the expected outcome of arbitration under any given state of the world. Define τ^{min}_θ and τ^{max}_θ as the payoff-equivalent action of D's and C's outside option, respectively, with

$$\begin{aligned} W^D\left(\tau^{min}_\theta;\theta\right) &\equiv W^D_A(\theta), \\ W^C\left(\tau^{max}_\theta\right) &\equiv W^C_A(\theta). \end{aligned}$$

[19] For example, non-settlement in a trade dispute may eventually lead to renegotiation of the WTO's arbitration between disputing parties, as analyzed by Beshkar (2016).

Lemma 2 implies that $\tau_\theta^{max} > \tau_\theta^{min}$, generating $\left[\tau_\theta^{min}, \tau_\theta^{max}\right]$ to be the core of the bargaining game if the state of the world is $\theta \in \{l, h\}$.
As demonstrated in the following analysis, despite the existence of these pre-arbitration negotiation deals that Pareto-dominate the expected non-settlement outcome, the equilibrium entails a positive probability of non-settlement due to the existence of asymmetric information.

3 Settlement Bargaining with an Imperfect Private Signal

In this section we analyze the impact of an imperfect private signal on the outcome of pre-arbitration settlement bargaining between C and D. To this end, we use a signaling model in which D's settlement proposal signals its type. The sequence of events is as follows:

Sequence of Events

1. State of the world, θ, is realized and observed privately by D.
2. C receives an imperfect private signal, θ^C.[20]
3. D proposes an action pair, $\tau^S \in R^+$, for settlement.
4. C either accepts τ^S with τ^S being implemented, or rejects τ^S and the dispute escalates to arbitration.

A strategy for D_θ is a function $\alpha_\theta(\tau^S)$ that specifies the probability that an action $\tau^S \in \mathbb{R}^+$ be proposed for settlement. A strategy for C_{θ^C} is a function $\beta_{\theta^C}(\tau^S) \in [0, 1]$ that specifies the probability that the complainant rejects D's proposal.

In a Perfect Bayesian Equilibrium (PBE), an initial equilibrium concept that we employ, the parties form consistent beliefs about each other's types.[21] Therefore, from the perspective of D of a type $\theta \in \{l, h\}$ who does not know the realized value of C's private signal, the likelihood that its settlement proposal, τ^S, will be

[20]While we analyze the case in which D has no access to this information (private imperfect signal) in this section, the following section considers the case in which D does have access to this information, namely, the public imperfect signal case.

[21]Because multiple PBEs arise as in other signal signaling games, we adopt *Universal Divinity* refinement of Banks and Sobel (1987) as discussed in Lemma 4 below.

rejected is given by $\gamma\beta_\theta\left(\tau^S\right)+(1-\gamma)\,\beta_{\theta^*}\left(\tau^S\right)$ with $\theta^*\neq\theta\in\{l,h\}$. Hence, the expected payoff of D_θ from proposing τ^S can be written as

$$\begin{gathered}\left(\gamma\beta_\theta\left(\tau^S\right)+(1-\gamma)\,\beta_{\theta^*}\left(\tau^S\right)\right)W_A^D(\theta)+\\ \left[1-\left(\gamma\beta_\theta\left(\tau^S\right)+(1-\gamma)\,\beta_{\theta^*}\left(\tau^S\right)\right)\right]W^D\left(\tau^S;\theta\right).\end{gathered}$$

Note that if the defendant offers $\tau^S=\tau_l^{max}$, it will be accepted by the complainant. This is because such an offer makes the complainant indifferent between arbitration and settlement if $\theta=l$, while the complainant will strictly prefer such an offer to arbitration if $\theta=h$. Therefore, if D_l chooses to 'separate' herself from D_h, its optimal strategy will be to offer $\tau^S=\tau_l^{max}$. Hence, in a (partially-)separating equilibrium in which D_l is at least indifferent between proposing τ_l^{max} and mimicking D_h, any proposal, τ^S, by D_h must satisfy[22]

$$\begin{aligned}W^D\left(\tau_l^{max};l\right)\geq&\left[\gamma\beta_l\left(\tau^S\right)+(1-\gamma)\beta_h\left(\tau^S\right)\right]W_A^D(l)+\\&\left[1-\gamma\beta_l\left(\tau^S\right)-(1-\gamma)\beta_h\left(\tau^S\right)\right]W^D\left(\tau^S;l\right).\end{aligned}\tag{3}$$

Among all τ^Ss that satisfy D_l's incentive compatibility constraint (3) for a separating equilibrium, D_h will choose one that maximizes its expected payoff, namely

$$\begin{aligned}\max_{\tau^S}&\left\{\left[\gamma\beta_h\left(\tau^S\right)+(1-\gamma)\,\beta_l\left(\tau^S\right)\right]W_A^D(h)+\right.\\&\left.\left[1-\left(\gamma\beta_h\left(\tau^S\right)+(1-\gamma)\,\beta_l\left(\tau^S\right)\right)\right]W^D\left(\tau^S;h\right)\right\}\\&s.t.(3),\end{aligned}\tag{4}$$

denoting the solution by τ_b.

In order to characterize the complainant's equilibrium strategy, we first show that

Lemma 3. *Under any PBE entailing τ^S with $\alpha_l\left(\tau^S\right)>0$ and $\alpha_h\left(\tau^S\right)>0$, we must have*

$$0<\beta_l\left(\tau^S\right)<1\Rightarrow\beta_h\left(\tau^S\right)=0,\ \textit{and}$$

$$\beta_h\left(\tau^S\right)>0\Longrightarrow\beta_l\left(\tau^S\right)=1.$$

[22] A (partially-)separating equilibrium is a PBE in which $\exists\tau^S$ with $\alpha_l\left(\tau^S\right)>0$ and $\alpha_h\left(\tau^S\right)=0$. From brevity, we refer a (partially-)separating equilibrium as a separating equilibrium, reserving "fully-separating" one for the equilibrium under which D_l and D_h choose separate actions, thus separate each other with probability 1.

To see this, note that $0 < \beta_l\left(\tau^S\right) < 1$ is true if and only if C_l is indifferent between arbitration and settlement at τ^S, which in turn implies that C_h must strictly prefer settlement since the expected outcome of arbitration is less favorable to the complainant if the state of the world is more likely to be $\theta = h$. The second condition above is true for a similar reason: if C_h is indifferent about arbitration and a given settlement offer, C_l will strictly prefer arbitration as his expected payoff with arbitration are better than that of C_h.[23]

If $0 < \beta_{\theta^C}\left(\tau^S\right) < 1$, the incentive compatibility condition of C_{θ^C} can be written as

$$W^C\left(\tau^S\right) = \Pr\left(\theta = l|\theta^C = j, \tau^S\right) W_A^C(l) + \Pr\left(\theta = h|\theta^C = j, \tau^S\right) W_A^C(h), \tag{5}$$

where,

$$\begin{aligned}\Pr\left(\theta = l|\theta^C = l, \tau^S\right) &= \frac{\Pr\left(\tau^S|\theta^C = l, \theta = l\right)\Pr\left(\theta^C = l, \theta = l\right)}{\Pr\left(\theta^C = l, \tau^S\right)} \\ &= \frac{\alpha_l\left(\tau^S\right)\gamma(1-\rho)}{\alpha_l\left(\tau^S\right)\gamma(1-\rho) + \alpha_h\left(\tau^S\right)(1-\gamma)\rho},\end{aligned}$$

$$\begin{aligned}\Pr\left(\theta = l|\theta^C = h, \tau^S\right) &= \frac{\Pr\left(\tau^S|\theta^C = h, \theta = l\right)\Pr\left(\theta^C = h, \theta = l\right)}{\Pr\left(\theta^C = h, \tau^S\right)} \\ &= \frac{\alpha_l\left(\tau^S\right)(1-\gamma)(1-\rho)}{\alpha_l\left(\tau^S\right)(1-\gamma)(1-\rho) + \alpha_h\left(\tau^S\right)\gamma\rho},\end{aligned}$$

for $\alpha_l\left(\tau^S\right) + \alpha_h\left(\tau^S\right) > 0$. To derive this expression for $\Pr\left(\theta = l|\theta^C = j, \tau^S\right)$, we use the fact that in equilibrium beliefs are consistent. Finally, $\beta_{\theta^C}\left(\tau^S\right) = 1$ $\left(\beta_{\theta^C}\left(\tau^S\right) = 0\right)$, only if C_{θ^C} at least weakly prefers arbitration (settlement).

Prior to a further analysis of PBEs, we introduce an additional assumption:

$$-\frac{\frac{\partial W^D(\tau;l)}{\partial \tau}}{\frac{\partial W^D(\tau;h)}{\partial \tau}} + \frac{W^D(\tau;l) - W_A^D(l)}{W^D(\tau;h) - W_A^D(h)} > 0 \tag{6}$$

[23]This logic for Lemma 3 does not apply for $\tau^S = \tau_h^{max}$. $\alpha_l(\tau_h^{max}) = 0$ under any PBE of our interest because $\alpha_l(\tau_h^{max}) > 0$ will induce $\beta_l\left(\tau^S\right) = \beta_h\left(\tau^S\right) = 1$ by making C's expected arbitration payoff be greater than $W^C\left(\tau_h^{max}\right) = W_A^C(h)$. If $\alpha_l(\tau_h^{max}) = 0$ then C's signal no longer carries any informational value on $\tau^S = \tau_h^{max}$, thus nullifying the above logic. However, Lemma 3 would still hold for $\tau^S = \tau_h^{max}$ under the refinement described in Lemma 4 below.

for $\tau \in \left[\tau_h^{min}, \tau_h^{max}\right]$, which simplifies the following analysis.[24] This inequality is a sufficient condition for the solution of the constrained maximization problem in (4), τ_b, to be τ_h^{max}, the highest settlement proposal among all PBEs.

This game has multiple Perfect Bayesian Equilibria as in other signaling games. The following lemma shows that the Divinity refinement narrows them down into one that is intuitively appealing:[25]

Lemma 4. *A PBE survives the Divinity refinement if and only if it is a separating equilibrium that maximizes D_h's expected payoff among all PBEs.*

For brevity, henceforth, we refer to the Divine PBE as the *equilibrium*. This equilibrium is an intuitively appealing one as the defendant is the party that proposes a take-it-or-leave-it offer so that she maximizes her expected payoff from such an offer. In addition, a high-type defendant can distinguish herself from a low-type defendant by utilizing the fact that D_h benefits more from proposing a higher-level action than D_l benefits, even though such an offer will trigger a costly arbitration with a higher probability. This results from the refinement of PBEs that defines the off-the-equilibrium beliefs of C.

The following proposition then characterizes the equilibrium of the game:

Proposition 1. *The equilibrium of pre-arbitration settlement bargaining with an imperfect private signal is a fully-separating equilibrium that is characterized by two action pairs, τ_l^{max} and τ_h^{max}, and two rejection probabilities, denoted by $\beta_{\theta^C}, \theta^C \in \{l, h\}$, such that*

(i) D_l always proposes τ_l^{max} for settlement, which will be accepted by the complainant;

(ii) D_h always proposes τ_h^{max}, which is rejected by C_{θ^C} with probability β_{θ^C} that are uniquely defined by (3) holding with equality and $\beta_h(1-\beta_l) = 0$ with $\beta_l > \beta_h$.

As discussed above, if D_l chooses to 'separate' herself from D_h, her optimal strategy will be to offer $\tau^S = \tau_l^{max}$ and settle for sure, thus generating (i): if D_l mimics D_h by proposing $\tau^S = \tau_h^{max}$ with a positive probability, then C will invoke arbitration for sure in the equilibrium, discouraging such behavior.

For D_h, proposing a higher level action increases her settlement payoff but it also increases the likelihood of costly arbitration being invoked. The inequality

[24] This inequality condition is satisfied as long as a possible positive level effect of a higher value of θ on D's payoff (i.e., $\partial W^D(\tau;\theta)/\partial\theta > 0$) is not large enough to reverse the inequality. If $\partial W^D(\tau;\theta)/\partial\theta \approx 0$, for example, then (6) holds with its first term being greater than -1 and its second term being also greater than 1. Even if $\partial W^D(\tau;\theta)/\partial\theta \gg 0$, it is difficult to have (6) be violated, as confirmed in the numerical example of Section 5.3.

[25] Although we adopt the *Universal Divinity* refinement, a stronger refinement concept than *Divinity*, we will refer a *Universally Divine* equilibrium as a *Divine* equilibrium for simplicity.

condition in (6) guarantees that the marginal benefit from raising its action level, thereby raising its settlement payoff is higher than the marginal cost from raising the arbitration likelihood for all $\tau^S \in \left[\tau_h^{min}, \tau_h^{max}\right]$, inducing D_h to propose τ_h^{max} for all values of $\gamma \in (0.5, 1)$.[26]

Our next proposition provides comparative statics results regarding the accuracy of the complainant's private signal.

Proposition 2. *An increase in the accuracy of the private signal, γ, will result in: (i) an increase in the probability of litigation against an imposter (i.e., $\gamma\beta_l\left(t_h^S\right) + (1-\gamma)\,\beta_h\left(t_h^S\right)$); (ii) a decrease in rejection probability (β_{θ^C}); (iii) an increase in the expected payoff of D_h; and (iv) no changes in the expected payoffs of D_l and C.*

By decreasing (increasing) the likelihood of arbitration against a truth-telling high-type defendant (an imposter), a more accurate private signal increases the total expected joint payoffs of the parties. However, due to the structure of bargaining in which the defendant proposes a take-it-or-leave-it offer, the high-type defendant, who succeeds in separating herself from the low-type one in the equilibrium, will extract all the extra rent generated by a more accurate signal.

4 Value of Second-Order Uncertainty in Dispute Settlement

So far, we have discussed settlement bargaining assuming that the complainant observes a signal of the defendant's type *privately*. The private nature of the signal generates second-order uncertainty in the relationship between the disputing parties: The defendant does not accurately know the complainant's belief about the defendant's type (or more precisely, the defendant's belief about the complainant's belief of her type is different from the complainant's belief). In order to obtain a deeper understanding of how private signals affect settlement negotiation, in this section we ask two questions.

First, would the efficiency of settlement bargaining improve if the second-order uncertainty was eliminated, i.e., if the signal observed by complainant was also observable to the defendant? We find that the answer to this question is negative. In fact, as we show in Subsection 4.1, a signal of the state of the world could improve the efficiency of negotiation *only if* it is privately observed by the complainant.

[26] If γ is sufficient large, we can also show that $\tau_b = \tau_h^{max}$ without (6).

To understand this result, suppose that the complaining party chooses to reject a settlement proposal with a higher probability if he receives a low signal. If the signal is informative, this strategy will induce an imposter to face a higher rate of rejection than a genuinely high-type defendant. This strategy could be part of an equilibrium if the induced probability of rejecting an imposter is high enough as to make a low-type defendant prefer, at least weakly, to report her type truthfully. With a public signal, however, it is impossible to have an equilibrium strategy that rejects the proposals of an imposter and a genuinely high-type defendant with different probabilities.[27]

The second question, discussed in Subsection 4.2, is whether or not an imperfect private signal continues to be of value if the bargaining game was a screening rather than a signaling game. In this regard, our first observation is that in a screening game, it does not matter whether the signal received by the uninformed party is private or public. The best offer strategy of the complainant depends only on the probability that the defendant is a high type conditional on his signal, which is unaffected by whether his signal is private or public. Then, the decision to settle is made by the informed party (i.e., the defendant) and, thus, whether the signal is private or public is inconsequential. In the screening game, we show that a signal is useful regardless of its privacy if and only if it is sufficiently informative.

4.1 Public Signal

We now analyze the pre-arbitration settlement game under the assumption that the signal observed by the complaining party is now *public* rather than private.[28] All other assumptions remain the same as the game under private signaling discussed in the previous section.

First consider separating PBEs. Assuming that C's signal is public rather than private changes the incentive compatibility constraint for the low-type D. Recall that when C's signal is private, D_l perceives the likelihood of arbitration given τ^S to be $\gamma\beta_l\left(\tau^S\right) + (1-\gamma)\,\beta_h\left(\tau^S\right)$. In contrast, when C's signal is public, D_l's perceived likelihood of arbitration given τ^S is $\beta_l\left(\tau^S\right)$ and $\beta_h\left(\tau^S\right)$ if $\theta^C = l$ and

[27] To be more precise, any pooling PBE that assigns different rejection rates depending on the signal, will not survive the Divinity refinement.

[28] While the term "public" is chosen to contrast it with "private," the signal being public does not necessarily means that such a signal is also automatically shared and utilized in the arbitration process if it is invoked. On the contrary, we continue to assume that the arbitrator ignores (i.e., is ignorant of) what happened in the settlement bargaining process.

$\theta^C = h$, respectively. Therefore, under the public signal, the incentive compatibility constraint for D_l (not to mimic D_h's settlement offer) given τ^S and θ^C can be written as

$$W^D\left(\tau_l^{\max}; l\right) \geq \beta_{\theta^C}\left(\tau^S\right) W_A^D(l) + \left(1 - \beta_{\theta^C}\left(\tau^S\right)\right) W^D\left(\tau^S; l\right) \text{ for } \theta^C = l, h. \quad (7)$$

The left-hand side of this condition is the welfare of D_l if it reveals her type by proposing $\tau_l^{\max}$. The right-hand side of this condition is D_l's expected welfare if she mimics D_h by proposing τ^S *given the realized public signal*, θ^C. Note that the public signal has no effect on this condition because these incentive compatibility constraints are identical regardless of realized value for θ^C. This implies that $\beta_l\left(\tau^S\right) = \beta_h\left(\tau^S\right)$ for separating PBEs with these constraints binding. Intuitively, in order to make D_l indifferent between proposing $\tau_l^{\max}$ and τ^S, for a given settlement proposal τ^S, C has to choose the same rate of rejecting τ^S *regardless of the realized public signal*. Therefore,

Lemma 5. *The separating PBEs of the settlement bargaining game are independent of the accuracy of the imperfect public signal.*

While separating PBEs could be a function of the imperfect public signal, note that the *accuracy* of the imperfect public signal is completely inconsequential for the equilibrium strategies.

Pooling Equilibria

We now investigate the impact of a public signal on pooling equilibria. With a public signal at the pre-arbitration stage, the common prior belief of the parties is updated. Let ρ^* denote the updated common belief of the parties about the likelihood of a high type after observing the public signal. Moreover, let $\tau_{pool}(\rho^*)$ denote a tariff pair that makes C indifferent between settlement and arbitration when all types of D pool, namely,

$$(1 - \rho^*) W_A^C(l) + \rho^* W_A^C(h) \equiv W^C\left(\tau_{pool}(\rho^*)\right). \quad (8)$$

$\tau_{pool}(\rho) < \tau_h^{\min}$ by our assumption about the quality of the court.[29] Moreover, as ρ^* increases, the expression on the left-hand side of (8) decreases, which implies

[29] By our assumption of an informative arbitrator, if the common prior belief about the likelihood of a high type is ρ, a fully-settling (i.e., parties always preferring settlement over arbitration) pooling offer cannot be an equilibrium. If the full-settlement pooling offer with $\tau \geq \tau_h^{min}$ is an equilibrium, then it implies that the arbitration cannot outperform (i.e., generate a higher expected joint payoff than from playing) such a full-settlement pooling equilibrium, which in turn contradicts Corollary 1.

that $\tau_{pool}(\rho^*)$ moves monotonically toward $\tau_h^{\max}$. If ρ^* is sufficiently high such that $\tau_{pool}(\rho^*) > \tau_h^{\min}$, then $\tau_{pool}(\rho^*)$ is part of a pooling PBE in which both types of D propose $\tau_{pool}(\rho^*)$ and C accepts with certainty. Therefore,

Lemma 6. *Let $\tau_{pool}(\rho^*)$ and $\hat{\rho}_1$ be defined by (8) and the following condition, respectively:*

$$(1-\hat{\rho}_1) W_A^C(l) + \hat{\rho}_1 W_A^C(h) \equiv W^C\left(\tau_h^{\min}\right).$$

If $\rho^* \geq \hat{\rho}_1$, then any $\tau^S \in \left(\tau_h^{\min}, \tau_{pool}(\rho^*)\right)$ constitutes a pooling PBE where both types of D propose τ^S for settlement and C accepts the proposal.

Therefore, although the accuracy of a public signal has no impact on the set of *separating* PBEs, a *pooling* PBE can arise with a sufficiently informative public signal. Nevertheless, as we show in the proof of the following Lemma in the Appendix, pooling PBEs are not Divine.[30] Moreover,

Lemma 7. *Under public signaling, a PBE is Divine iff it is a separating PBE that maximizes the expected payoff of D_h.*

The following proposition then characterizes the equilibrium of the game with an imperfect public signal:

Proposition 3. *The equilibrium of pre-arbitration settlement bargaining with an imperfect public signal is a fully-separating equilibrium that is characterized by two action pairs, τ_l^{max} and τ_h^{max}, and one rejection probability, denoted by β with $\beta_{\theta^C} = \beta$, $\theta^C \in \{l, h\}$, such that*

(i) D_l always proposes τ_l^{max} for settlement, which will be accepted by the complainant;

(ii) D_h always proposes τ_h^{max}, which is rejected by C with probability β that is uniquely defined by (7) holding with equality so that

$$\beta = \frac{W^D\left(\tau_h^{max}; l\right) - W^D\left(\tau_l^{\max}; l\right)}{W^D\left(\tau_h^{max}; l\right) - W_A^D(l)}.$$

[30] Any pooling settlement proposal entails τ^S that is strictly less than τ_h^{max} (because $\tau^S \geq \tau_h^{max}$ induces C to reject it for sure under a pooling equilibrium), and then D_h would have an incentive to deviate from such a pooling equilibrium by raising τ^S to a higher level, possibly to τ_h^{max}, and settle for sure: According to the Divinity refinement, C would believe that such an off-equilibrium proposal is made by D_h (not D_l), which in turn implies that C will accept the deviation proposal and settle for sure.

In contrast to the equilibrium with an imperfect private signal in which C's rejection probability changes in C's signal accuracy, the equilibrium strategy with an imperfect public signal stay the same regardless of C's signal accuracy. When $\gamma \to 0.5$, one can also easily show that the equilibrium strategy and payoffs with a public signal are (qualitatively) identical to those with a private signal.[31] Since we have shown in Proposition 2 that a private signal improves the expected joint payoff of the parties by raising the expected payoff of the high-type D, Proposition 3 together with Proposition 2 imply that

Theorem 1. *An informative signal about D's private information improves the expected payoffs of the parties in the signaling game iff it is privately observed by C.*

The following thought experiment provides some perspective on the result that a signal is useful only if it is private. Consider an action $\tau^S \in \left[\tau_h^{\min}, \tau_h^{\max}\right]$, and the equilibrium strategies that support it as a separating PBE under public and private signaling, respectively. Letting β denote the likelihood of arbitration if τ^S is proposed under public signaling, the incentive compatibility constraint for D_l is given by

$$W^D\left(\tau_l^{\max}; l\right) \geq \beta W_A^D(l) + (1-\beta) W^D\left(\tau^S; l\right).$$

Solving for β that makes D_l indifferent yields

$$\beta = \frac{W^D\left(\tau^S; l\right) - W^D\left(\tau_l^{\max}; l\right)}{W^D\left(\tau^S; l\right) - W_A^D(l)}.$$

The value of β indicates the likelihood that D_h, as well as untruthful D_l, will face arbitration in the equilibrium of a public signaling game.

In the case of private signaling, the probability of arbitration for an untruthful D_l is given by $\gamma\beta_l + (1-\gamma)\beta_h$, and the incentive compatibility constraint is given by

$$W^D\left(\tau_l^{\max}; l\right) \geq \left(\gamma\beta_l + (1-\gamma)\beta_h\right) W_A^D(l) + \left(1 - \left(\gamma\beta_l + (1-\gamma)\beta_h\right)\right) W^D\left(\tau^S; l\right).$$

[31] While $\beta_l > \beta_h$ in the equilibrium with a private signal and $\beta_l = \beta_h = \beta$ in the equilibrium with a private signal, $\gamma\beta_l + (1-\gamma)\beta_h$ in the equilibrium with a private signal is equal to β as discussed below. When $\gamma \to 0.5$, then $(1-\gamma)\beta_l + \gamma\beta_h \to \beta$, generating the same equilibrium arbitration probability in both cases.

Solving for $\gamma\beta_l + (1-\gamma)\beta_h$ yields

$$\begin{aligned}\gamma\beta_l + (1-\gamma)\beta_h &= \frac{W^D\left(\tau^S;l\right) - W^D\left(\tau_l^{\max};l\right)}{W^D\left(\tau^S;l\right) - W_A^D\left(l\right)} \\ &= \beta.\end{aligned}$$

Therefore, the incentive compatibility constraint implies the same likelihood of arbitration for the low type D under private and public signaling.

Now consider the likelihood of arbitration for D_h under private signaling, which is given by $\gamma\beta_h + (1-\gamma)\beta_l$. For $\gamma > \frac{1}{2}$, we have

$$\gamma\beta_h + (1-\gamma)\beta_l < \gamma\beta_l + (1-\gamma)\beta_h,$$

which implies that D_h faces a lower likelihood of arbitration under private signaling than under public signaling.[32]

This comparison clarifies the source of efficiency improvement under private signaling: when the signal is unobservable to D, C can condition its arbitration strategy on the private signal and litigate different types with different probabilities. In particular, under private signaling, D_h is less likely to be litigated than untruthful D_l.

4.2 A Screening Game of Settlement Bargaining

In this section, we reformulate our settlement bargaining game as a *screening* game in which the uninformed party (i.e., the complainant) makes a settlement proposal and the informed party (i.e., the defendant) decides whether to accept the proposal and settle, or reject the proposal and demand arbitration. We assume the following sequence of events. First, C receives a signal of D's type and proposes an action for settlement. D can either drop the case and adopt the status quo, $\tau_l^{\min}$, accept the proposal and settle, or demand arbitration..

D will accept the proposal if $W\left(\tau^S;\theta\right) \geq W_A^D\left(\theta\right)$. D will drop the case and implement the status quo, $\tau_l^{\min}$, iff $\theta = l$ and $W^D\left(\tau^S;l\right) < W_A^D\left(l\right)$. Finally, D will demand arbitration if D's expected payoff under arbitration is greater than the payoff under the status quo and the proposed settlement, i.e., $W_A^D\left(\theta\right) > W^D\left(\tau_l^{\min};\theta\right)$ and $W_A^D\left(\theta\right) > W^D\left(\tau^S;\theta\right)$.

[32]Recall that $0 < \beta_l\left(\tau^S\right) < 1 \Rightarrow \beta_h\left(\tau^S\right) = 0$ and $\beta_h\left(\tau^S\right) > 0 \Longrightarrow \beta_l\left(\tau^S\right) = 1$ from lemma 3, which in turn implies $\beta_l\left(\tau^S\right) > \beta_h\left(\tau^S\right)$. The above inequality results from this inequality together with $\gamma > \frac{1}{2}$.

Proposition 4. *Under the screening game: i) For* $\Pr(\theta = h|\theta^C) > \frac{W^C(\tau_l^{\min}) - W^C(\tau_h^{\min})}{W^C(\tau_l^{\min}) - W_A^C(h)}$, *the unique PBE is a pooling equilibrium in which C proposes* $\tau_h^{\min}$ *for settlement and both types of D will accept this proposal.*

ii) For $\Pr(\theta = h|\theta^C) < \frac{W^C(\tau_l^{\min}) - W^C(\tau_h^{\min})}{W^C(\tau_l^{\min}) - W_A^C(h)}$, *the unique PBE is a separating equilibrium in which C proposes* $t_l^{\min}$ *for settlement, which will be accepted (rejected) by* D_l *(*D_h*).*

Proposition 4 does not depend on whether C's signal is private or public. Therefore, in a screening game, the second-order uncertainty is immaterial. The decision to settle is not made by the uninformed party (i.e., complainant), which in turn makes second-order uncertainty about the uninformed party's belief irrelevant.

5 Settlement over Action Combinations: Dispute Settlement of the WTO

In this section, we extend our simple model by allowing not only a defendant but also a complainant to adjust his own action that may benefit himself but hurts the other party. For example, each country in trade disputes can adjust her own import tariff level and both countries' tariffs can be a subject of bargaining. With this extension, we can apply our settlement bargaining model with second order uncertainty to the analysis of WTO trade disputes and settlements.

The obligations of an importing country under the WTO are *contingent* on its domestic political economy conditions, which are likely to be the private information of the importing government.[33] Other governments, however, could also conduct their own investigations and receive informative signals about the importing country's political economy conditions. These signals, which are potentially the private information of the investigating governments, creates second-

[33] Various types of contingent protection are allowed under the WTO regime. They can be safeguard measures, anti-dumping measures, countervailing duties against export subsidies, or even general exceptions to General Agreement of Tariffs and Trade (GATT, the agreement prior to the WTO, that remains a central part of the WTO regime) obligations. These general exceptions are given based on two articles, Article XX to protect public morals, human, animal, or plant life or health, international intellectual property rights, and etc., and Article XXI to protect national security. Sykes (2016) provides a comprehensive description of these contingent protection measures and other aspects of legal obligations of WTO.

order uncertainty in the pre-arbitration bargaining game.[34] If disputing parties fail to settle, then Dispute Settlement Body of the WTO provides its ruling on the disputed case and possibly authorizes a complainant's compensatory/retaliatory protection against a defendant who refuses to follow its ruling.

With regard to the significance of settlement through pre-arbitration bargaining, an WTO's official website states "..., the point is not to pass judgment. The priority is to settle disputes, through consultations if possible. By January 2008, only about 136 of the nearly 369 cases had reached the full panel process. Most of the rest have either been notified as settled 'out of court' or remain in a prolonged consultation phase - some since 1995." Of the economic importance of trade disputes filed to the WTO dispute settlement procedure, Bown and Reynolds (2015) find that the value of imported goods subject to the WTO disputes from 1995 to 2011 is almost $1 trillion, an average of $55 billion per year, or equivalently about 0.5 percent of world imports in 2011. Given that only a small portion of trade disputes ends up being filed to WTO, a much bigger percentage of world imports must be under trade disputes, implying that trade disputes and settlements affect a sizable portion of the world trade.

A defendant government typically commits herself to a specific level of import protection through a domestic process of determining the legitimacy of such protection prior to a complainant's litigation decision. Thus, our signaling model in Section 3 (rather than the screening model in Section 4.2) is appropriate to study trade disputes and settlements in the WTO. As discussed in Section 5.1 on a WTO dispute case, trade disputes may entail settlement over both parties' protection measures. This necessitates extending our simple model by allowing not only a defendant but also a complainant to adjust his own action as in Section 5.2. Section 5.3 then provides an numerical example with a trade model, comparing the two-action-variable case with the one-action-variable one.

5.1 A WTO Dispute Case: Slovakia — Safeguard Measure on Imports of Sugar

For concrete understanding of trade disputes and settlements, we provide a detailed discussion of the WTO dispute case number 235.[35] Poland filed a consultation request on July 11, 2001, claiming that Slovakia had imposed a safeguard

[34] See footnote 7 for the discussion of Best Information Available cases of the U.S. anti-dumping investigations as an example.

[35] For some empirical analyses of the WTO dispute settlement process, see Ahn et al. 2014; Beshkar and Majbouri 2019; Bown 2004; Bown and Reynolds 2017; Kuenzel 2017.

measure on imports of sugar in a manner inconsistent with the obligations under the Agreement on Safeguards, with the following statement as one of her key claims in the consultation request document:

"No document presented, including notifications to the SG Committee under Article 12, contained analyses on *causal link between increased imports and serious injury to the domestic industry* or factor other than imports which might have caused injury."

A safeguard measure is justifiable if *a causal link between increased imports and serious injury to the domestic industry* exists, upon which disputing parties may disagree.[36] With regard to the strength of this case, the defending government is likely to have some private information, such as assessment of the degree of damages to her domestic firms' profitability caused by increased imports. A complaining government may also have an imperfect and private signal of the causal link that is provided by her own domestic exporting firms to persuade her to file a dispute case to the WTO, as discussed in footnote 7. The outcome of this dispute was a mutually agreed solution as follows:

"On 11 January 2002, the parties notified the arbitrator that they have reached a mutually agreed solution within the meaning of Article 3.6 of the DSU. Accordingly, Slovakia agreed to a progressive increase of the level of its quota of imports of sugar from Poland between 2002 and 2004, and Poland agreed to remove its quantitative restriction on imports of butter and margarine. Both parties agreed to implement the above by 1 January 2002."

This report demonstrates that a settlement is achieved by exchanging desired trade policies among disputing parties. We may interpret this settlement as an exchange of temporary protection policies.

5.2 Settlement over Action Combinations

In this subsection, we extend our analysis by allowing not only a defendant but also a complainant to adjust his own action variable and to settle on a combination of their actions. Thus, there are two parties, (D)efendant and (C)omplainant, each with a corresponding action variable $\tau_D, \tau_C \in R^+$. The parties' payoffs are given by $W^D(\tau_D, \tau_C; \theta)$ and $W^C(\tau_D, \tau_C)$, respectively, where, θ represents the state of the world that affects D's payoff (and only D's). In the context of trade disputes, θ

[36] Beshkar and Bond (2016) provide a literature review of studies on safeguard measures in the WTO and other trade agreements.

may represent the importing country's political economy condition that affects her (D's) desirability to raise her import tariff rate, τ_D, with τ_C denoting C's import tariff rate.[37] With regard to how this addition of an action variable affects payoffs, we maintain the same assumptions 1-4 on the derivatives of τ_D in Section 2, and keep the assumptions 1, 2, and 4 on the derivative of τ_C on the corresponding payoffs with

$$\frac{\partial^2 W^C}{\partial \tau_C^2} < 0, \left.\frac{\partial W^C}{\partial \tau_C}\right|_{\tau=0} > 0,$$

$$\frac{\partial W^D}{\partial \tau_C} < 0,$$

$$\frac{\partial^2 \left(W^D + W^C\right)}{\partial \tau_C^2} < 0,$$

plus the following additional assumption for $i, j = \{C, D\}$:

5. Cross partial derivatives of payoffs in actions are zero:

$$\frac{\partial^2 W^i}{\partial \tau_i \partial \tau_{j \neq i}} = 0.$$

Then, $\tau_D^N(\theta) > \tau_D^E(\theta)$ again, and $\tau_C^N > \tau_C^E$. We also keep the assumption that the state of the world can take one of two levels: h and l with $h > l$, thus $\tau_D^N(h) > \tau_D^N(l)$ and $\tau_D^E(h) > \tau_D^E(l)$. Finally, we continue to assume that $\tau_D^N(l) > \tau_D^E(h)$.

With regard to the information structure and the arbitration, we maintain the same assumptions as the ones in Section 2. Then, the results in Lemma 1, Corollary 1, and Lemma 2 continue to hold with

$$\begin{gathered} \tau_C^A(h) = \tau_C^A(l) = \tau_C^E, \\ \tau_D^E(l) < \tau_D^A(l) < \tau_D^A(h) < \tau_D^E(h), \\ \forall t, E_\theta\left[W_A^D(\theta) + W_A^C(\theta)\right] > E_\theta\left[W^D(t;\theta) + W^C(t)\right], \\ \forall \theta \in \{l, h\}, \exists t : W^D(t, \theta) > W_A^D(\theta) \text{ and } W^C(t, \theta) > W_A^C(\theta), \end{gathered} \tag{9}$$

where $t \equiv (\tau_D, \tau_C) \in R^{++}$. Because $\partial^2 W^i / \partial \tau_i \partial \tau_{j \neq i} = 0$ and the state of world only influences how an increase in τ_D affects D's payoff with $\partial^2 W^D / \partial \tau_D \partial \theta > 0$, the arbitrator who tries to maximize the joint payoff should set $\tau_C^A(h) = \tau_C^A(l) = \tau_C^E$:

[37] Assumptions 1-5 specified below are standard in the trade policy literature.

the arbitrator knows for sure what is τ_C that maximizes the joint payoff. Then, for other results in (9), one can apply the same logic/proofs as the ones for the corresponding results in Section 2.[38]

To analyze the pre-arbitration bargaining game, once again we describe each party's *outside* option based on the expected outcome of arbitration under a given state of the world. Let P_θ denote the set of Pareto efficient action pairs under θ and define $t_\theta^D \equiv \left(\tau_D^\theta, \tau_C^\theta\right), t_\theta^C \equiv \left(\tau_D^\theta, \tau_C^\theta\right) \in P_\theta$ as the payoff-equivalent action pair of D and C's outside option, respectively, with

$$W^D\left(t_\theta^D;\theta\right) \equiv W_A^D\left(\theta\right),$$
$$W^C\left(t_\theta^C\right) \equiv W_A^C\left(\theta\right).$$

These action pairs are depicted in Figure 1, in which $T^C(t)$ and $T^{D_\theta}(t)$ denote the indifference curves of C and D_θ that go through t. The points between t_l^D and t_l^C on the contract curve P_l are the core of the bargaining game if the state of the world is $\theta = l$. Similarly, if the state of the world is $\theta = h$, the bargaining core consists of points between t_h^D and t_h^C on P_h. As we move from t_θ^D towards t_θ^C on P_θ, the payoffs of D (C) increases (decreases).

As discussed earlier, we will focus on the signaling model, having the sequence of events be defined in the same way as the one in Section 3 with one modification: D proposes an action pair, $t^S \in R^{++}$ (instead of an action, $\tau^S \in R^+$), for settlement in the third step of the sequence. Once again, a strategy for D_θ is a function $\alpha_\theta(t^S)$ that specifies the probability that an action pair $t^S \in \mathbb{R}^{++}$ be proposed for settlement. A strategy for $C_{\theta C}$ is a function $\beta_{\theta C}(t^S) \in [0,1]$ that specifies the probability that the complainant rejects D's proposal, t^S.

If D_l chooses to 'separate' herself from D_h, then her optimal strategy will be to offer $t^S = t_l^C$ as it maximizes D_l's payoff without provoking $C's$ arbitration request. Hence, in a (partially-)separating equilibrium in which D_l is at least indifferent between proposing t_l^C and mimicking D_h's proposal, D_h will try to maximize her expected payoff by proposing t^S subject to D_l's incentive compatibility

[38]For the inequality of the second line in (9), the arbitrated actions on D's action variable conditional on the arbitrator's information will be identical to those characterized in Lemma 1. This is because $\partial^2 W^i/\partial\tau_i\partial\tau_{j\neq i} = 0$ which makes the arbitrator's maximization problem with regard to the choice of τ_D be identical to the one for the one variable case. Then, the inequality of the third line in (9) follows as a corollary to this result as Corollary 1 follows from Lemma 1. To prove the last inequality in (9), one can apply the same proof as the one for Lemma 2, having τ_C set at τ_C^E.

constraint:

$$\begin{array}{c} \max_{t^S} \; \left\{ \left(\gamma\beta_h \left(t^S\right) + \left(1-\gamma\right) \beta_l \left(t^S\right) \right) W_A^D \left(h\right) + \right. \\ \left. \left[1 - \left(\gamma\beta_h \left(t^S\right) + \left(1-\gamma\right) \beta_l \left(t^S\right) \right) \right] W^D \left(t^S; h\right) \right\} \\ s.t. \\ W^D \left(t_l^C; l\right) \geq \left\{ \left[\gamma\beta_l \left(t^S\right) + (1-\gamma)\beta_h \left(t^S\right) \right] W_A^D \left(l\right) + \right. \\ \left. \left[1 - \gamma\beta_l \left(t^S\right) - (1-\gamma)\beta_h \left(t^S\right) \right] W^D \left(t^S; l\right) \right\} \end{array} \tag{10}$$

With regard to the characterization of PBEs, we can obtain the same results as in Lemma 3 and Lemma 4.[39] Thus, we focus on a separating equilibrium that maximizes D_h's expected payoff among all PBEs, obtaining the following proposition:

Proposition 5. *The equilibrium of the pre-arbitration settlement bargaining is a separating equilibrium that is characterized by two action pairs, t_l^C and $t_h^S \equiv \left(\tau_D^S\left(h\right), \tau_C^S\left(h\right)\right)$, and two rejection probabilities, denoted by β_{θ^C}, with $\theta^C \in \{l, h\}$, such that*

(i) D_l always proposes t_l^C for settlement, which will be accepted by the complainant;

(ii) D_h always proposes t_h^S which is rejected by C_{θ^C} with probability β_{θ^C}:

$t_h^S \in T^C\left(t_h^C\right)$ is the solution of (10), with $t_h^S \gg t_h^C$ and *β_{θ^C} being uniquely defined by D_l's incentive constraint in (10) holding with equality.*

An interesting feature of the equilibrium described in Proposition 5 is that D_h's proposed settlement, t_h^S, is Pareto inefficient, i.e. $t_h^S \notin P_h$. This result can be explained by two steps. First, D_h proposes $t^S \in T^C\left(t_h^C\right)$ because she would strictly prefer a point on $T^C\left(t_h^C\right)$ to any feasible settlement proposal, including a Pareto-efficient proposal like $t^/$ in Figure 1. Note that any proposal on $T^{D_l}\left(t^/\right)$ generates the same settlement payoff to D_l, thus associated with the same rejection probability to satisfy the incentive constraint in (10) as long as it belongs to the lens-shaped area between $T^{D_h}\left(t_h^D\right)$ and $T^C\left(t_h^C\right)$.[40] Among such proposals with the same rejection probability, D_h would prefer $t^S = t_h^S \in T^C\left(t_h^C\right)$ as it generates her the highest

[39] In contrast to Section 3 in which we introduce the condition in (6) to prove Lemma 4, we do not need any further condition in obtaining the same result when parties bargain over a combination of actions. As discussed in connection with Proposition 5 below, the added flexibility in settlement offering created by an additional action variable enables D_h to raise her settlement payoff without increasing the rejection probability. This makes the Divine off-the-equilibrium belief on D's type who offers $t^S \in T\left(t_h^C\right)$ to be D_h rather than D_l without any additional condition, such as (6). The basic logic for proving these results and the following Proposition 5 of the two action variable case is very similar to the one for proving corresponding results of the one action variable case. Thus, we refer formal proofs of these results to Beshkar and Park (2017).

[40] If it does not belong to this lens-shaped area, then either D or C prefer arbitration over it, thus leading to arbitration with probability 1.

settlement payoff. Second, $t_h^S \gg t_h^C \in P_h$. While $D_h's$ settlement payoff increases as she decreases t^S along $T^C\left(t_h^C\right)$ toward t_h^C, it will also strictly increase the probability of rejection that will lead to costly arbitration. As the marginal settlement payoff gain from decreasing t^S toward t_h^C diminishes to zero at $t^S = t_h^C$, it is never optimal to lower the settlement offer to t_h^C. Intuitively, by proposing this inefficient action pair, D_h makes it more costly for D_l to mimic a high type, thereby inducing the complainant to accept her settlement offer with a greater likelihood.

We can compare the one-action variable case in Section 3 with the two-action variable case as follows. D with two action variables could have proposed an action pair that are equivalent to the action proposed under the equilibrium with one action variable: $\left(\tau_l^{max}, \tau_C^E\right)$ by D_l and $\left(\tau_h^{max}, \tau_C^E\right)$ by D_h, which are denoted by t_l^{max} and t_h^{max}, respectively in Figure 2. In these one-variable equilibrium equivalent proposals, first note that τ_C is set at the efficient level, τ_C^E, so that the joint payoff is already maximized with regard to the choice over τ_C.[41] Proposing such an action pair would have resulted in the equilibrium that is practically identical to the one described in Proposition 1 in Section 3.[42] The equilibrium characterization in Proposition 5 then reveals that D_h can benefit from having one more action variable as a part of her settlement offer. As shown in Figure 2, D_h can maintain her settlement payoff by proposing A instead of proposing t_h^{max}. The benefit from proposing A instead of t_h^{max} comes from reducing C's rejection probability by making mimicking D_h's proposal less attractive to D_l: D_l's settlement payoff on $T^{D_l}(A)$ is lower than the one on $T^{D_l}\left(t_h^{max}\right)$. D_h may raise the expected payoff from his settlement offer even further by proposing $t_h^S\ (\neq A)$ as the marginal cost associated with proposing $t^S \in T^C\left(t_h^C\right)$ closer to t_h^C (an increase the rejection probability) can be different from the marginal benefit of such an action (an increase in the settlement payoff) at $t^S = A$.

D_l can also benefit from having one more action variable in her settlement offer, as shown in Figure 2: D_l's settlement payoff on $T^{D_l}\left(t_l^C\right)$ is higher than the one on $T^{D_l}\left(t_l^{max}\right)$. The expected payoff of C remains unchanged at $\rho W_A^C(h) + (1-\rho) W_A^C(l)$. Based on these observations, we can state the following corollary:

Corollary 2. *Allowing a defendant to include a complainant's action as a part of her settlement proposal (in addition to her own action) raises the expected joint payoff of the*

[41] In the absence of random shocks that affect the desirable level of τ_C, it is also reasonable to assume that the parties would agree to set $\tau_C = \tau_C^E$.

[42] To be more precise, the total payoff of each party is affected in a lump-sum way through the action, τ_C^E. However, once such a level-influence on each party's payoff is controlled away, characteristics of the resulting equilibrium will be identical to the ones characterized in Section 3.

pre-arbitration bargaining game.

With regard to the WTO dispute settlement, the above results have the following implications. First, $t_h^S \equiv \left(\tau_D^S(h), \tau_C^S(h)\right) \in T^C\left(t_h^C\right)$ with $t_h^S \gg t_h^C$ in Proposition 5 may explain the pro-trade bias in the WTO DSB rulings, which Legal scholars identified in their analysis.[43] Recall that $\tau_D^A(l) < \tau_D^A(h) < \tau_D^E(h)$ from (9) and D_h may commit to $\tau_D^S(h) > \tau_D^E(h)$ prior to arbitration, as shown in Figure 1. This implies that the arbitrator will require D_h to reduce his protection level even under a ruling that favors D based on her high-state-of-the-world judgment. This results from D_h's excessive protection level to mitigate D_l's mimicking incentive, which does not arise in the one variable case with $\tau^S = \tau_h^{max} < \tau^A(h)$.

The second implication is a normative one regarding the design of WTO pre-arbitration settlement. According to Corollary 2, it is desirable to allow disputants to negotiate a settlement deal not only on disputed trade policy measures of a defendant but also on potentially non-disputed trade policy measures of a complainant. As quoted above, the WTO's position on pre-arbitration settlement is summarized by the following statement, "The priority is to settle disputes, through consultations if possible." In comparison to the WTO arbitration process that focuses on disputed trade policy measures of a defendant, the WTO places practically no structure on the way that countries may settle their dispute through consultations. The WTO dispute settlement may involve an exchange of protection among disputing parties, as in the WTO dispute case number 235 discussed above.

Our next two propositions provides comparative statics results regarding the accuracy of the complainant's private signal.

Proposition 6. *An increase in the accuracy of the private signal, γ, will result in: (i) an increase in the probability of litigation against an imposter (i.e., $\gamma\beta_l\left(t_h^S\right) + (1-\gamma)\beta_h\left(t_h^S\right)$); (ii) an increase in the expected payoff of D_h; and (iii) no changes in the expected payoffs of D_l and C.*

By increasing the likelihood of litigation against an imposter, thus discouraging D_l's incentive to mimic D_h's proposal more effectively, a more accurate private signal raises the expected joint payoff associated with pre-arbitration settlement bargaining. Once again, due to the structure of bargaining, D_h will extract all the extra rent generated by a more accurate signal.

[43] For example, Sykes (2003) points out that the DSB has always ruled against the defending party in litigation regarding safeguard measures.

The following proposition complements Proposition 6 by describing the effect of a higher signal accuracy on D_h's equilibrium settlement proposal ($t_h^S \equiv \left(\tau_D^S(h), \tau_C^S(h)\right)$) and C's rejection probability (β_l and β_h).

Proposition 7. *There are three thresholds of $\gamma < 1$, $\gamma^I < \gamma^{II} < \gamma^{III}$ such that*

(i) If $\gamma < \gamma^I$, t_h^S stays constant at a level denoted by t_h^b, $\beta_l = 1$ and $\beta_h > 0$, with $\frac{\partial \beta_h}{\partial \gamma} < 0$;

(ii) If $\gamma^I \leq \gamma \leq \gamma^{II}$, $\beta_l = 1$ and $\beta_h = 0$, with $\frac{\partial \tau_D^S(h)}{\partial \gamma} < 0$ and $\frac{\partial \tau_C^S(h)}{\partial \gamma} < 0$;

(iii) If $\gamma > \gamma^{II}$, $\beta_l \in (0, 1]$ and $\beta_h = 0$, with $\frac{\partial \tau_D^S(h)}{\partial \gamma} < 0$ and $\frac{\partial \tau_C^S(h)}{\partial \gamma} < 0$;

(iv)If $\gamma \geq \gamma^{III}$,$\beta_l \in (0, 1)$, $\beta_h = 0$, and $\lim_{\gamma \to 1} t_h^S = t_h^C$.

In contrast to the one-variable case in which the rejection probability (β_l or β_h) strictly decreases with the settlement proposal being fixed (at $\tau^S = \tau_h^{max}$) when C's signal accuracy improves, the rejection probability and the settlement proposal change in a more complex manner when D's settlement offer includes two variables, τ_D and τ_C. This result comes from not only the rejection probability but also the settlement proposal adjusting in response to a rise in γ. The following representation of the first order condition associated with the maximization problem in (10) facilitates understanding the results in Proposition 7:

$$\begin{aligned} MB(t) \equiv \left[(1-\gamma)\tfrac{\partial \beta_l(t)}{\partial t} + \gamma \tfrac{\partial \beta_h(t)}{\partial t}\right]\left[W^D(t; h) - W_A^D(h)\right] = \\ MC(t) \equiv \left[1 - ((1-\gamma)\beta_l + \gamma\beta_h)\right]\tfrac{\partial W^D(t;h)}{\partial t} \end{aligned} \tag{11}$$

$MB(t)$ represents D_h's marginal benefit of proposing $t_h^S \in T^C\left(t_h^C\right)$ further away from t_h^C, reducing the rejection probability to satisfy D_l's incentive constraint in (10) as a result, and $MC(t)$ denotes D_h's marginal cost of proposing t_h^S further away from t_h^C, decreasing his settlement payoff as a result. Although the maximization problem in (10) involves choosing over a pair of tariffs (τ_D, τ_C) instead a single variable, we can (and will) treat $t \equiv (\tau_D, \tau_C) \in T^C(t_h^C)$ as if it is a single variable in the use of derivatives for solving the maximization problem.[44]

When the signal's accuracy is not sufficiently high ($\gamma < \gamma^I$), C needs to reject t_h^S with a positive probability even when he receives a high signal (i.e., $\beta_h > 0$

[44]Given that any change in the choice of t occurs along C's indifference curve, $T^C(t_h^C)$, we can let ∂t represent a change in t such that $\partial \tau_D = \partial t$ with $\partial \tau_C$ being defined by $(\tau_D + \partial \tau_D, \tau_C + \partial \tau_C) \in T^C(t_h^C)$. We can do this because we focus on the range of $T^C(t_h^C)$ that is positively sloped in the space of (τ_D, τ_C) without loss of generality.

with $\beta_l = 1$) in order to deter D_l from proposing t_h^S. With respect to an (small) improvement in the signal's accuracy, only β_h (with $\beta_l = 1$) needs to adjust to satisfy D_l's incentive constraint in (10), thus $\partial\beta_l(t)/\partial t = 0$ as a result. Then, an increase in γ raises $MB(t)$ through increasing the weight on the benefit from lowering the rejection probability ($\partial\beta_h/\partial t < 0$), but it also raises $MC(t)$ by the same magnitude through increasing the weight on the cost from a lower settlement payoff ($\partial W^D(t;h)/t < 0$): a higher γ implies a higher settlement likelihood with $\beta_l = 1 > \beta_h$. $MB(t;\gamma)$ and $MC(t;\gamma)$ are shown for different values of γ in Figure 3, having bold lines represent them for $\beta_l = 1$ and $\beta_h > 0$, and thin lines denote them for $\beta_l < 1$ and $\beta_h = 0$. The bold lined $MB(t;\gamma)$ and $MC(t;\gamma)$ crossed at $t_h^S = t_h^b$ for $\gamma = 0.5$ shift up by the same amount when γ rises to γ^I (or to any $\gamma \leq \gamma^I$), implying that t_h^S remains constant at t_h^b in Figure 3. With t_h^S being fixed at t_h^b, then an improvement in signal's accuracy enables C to deter D_l's imposter behavior with a lower rejection probability ($\partial\beta_h/\partial\gamma < 0$) as long as $\gamma < \gamma^I$.

If the accuracy of C's signal reaches a critical level, denoted by γ^I, then C can restrain D_l from proposing $t_h^S = t_h^b$ by rejecting the settlement proposal only when he receives a low signal with $\beta_h = 0$ and $\beta_l = 1$. In response to a further increase in γ, β_l remains constant at one if the signal accuracy is below a certain level ($\gamma \leq \gamma^{II}$). Note that an increase in γ raises $MB(t)$ if $\beta_h > 0$ as discussed above, but it lowers $MB(t)$ if $\beta_h = 0$: an increase in γ reduces the the weight on the benefit from lowering the rejection probability (β_l) with $\beta_l > \beta_h = 0$. For a given value of γ, this implies that $MB(t)$ suddenly jumps down once t_h^S is chosen sufficiently away from t_h^C so that $\beta_h > 0$ is no longer required to deter D_l from proposing t_h^S. Such a critical level of t_h^S at which $MB(t)$ jumps down is denoted by $\hat{t}_h^S(\gamma)$ in Figure 3. When γ increases to γ^I, t_h^S remains constant at t_h^b because $\hat{t}_h^S(\gamma^I) = t_h^b$ by definition of γ^I. An increase in γ beyond γ^I lowers $\hat{t}_h^S(\gamma)$ further, yielding $\hat{t}_h^S(\gamma) \ll t_h^b$ for $\gamma > \gamma^I$. Then, $MC(t) > MB(t)$ for $t_h^S \gg \hat{t}_h^S(\gamma)$ and $MC(t) < MB(t)$ for $t_h^S \ll \hat{t}_h^S(\gamma)$, implying $t_h^S = \hat{t}_h^S(\gamma) < t_h^b$ for $\gamma \in (\gamma^I, \gamma^{II}]$. Finally, note that $\hat{t}_h^S(\gamma)$ is the proposal for which C can restrain D_l from proposing $t_h^S = \hat{t}_h^S(\gamma)$ with $\beta_h = 0$ and $\beta_l = 1$, and $\hat{\partial t}_h^S(\gamma)/\partial\gamma < 0$. Thus, an improved signal accuracy in this range of γ will only induce D_h to propose t_h^S closer to t_h^C without affecting the rejection probability.

Once the accuracy of C's signal increases beyond another critical level, denoted by γ^{II}, then it is possible to have $MB(t) = MC(t)$ for $t_h^S \gg \hat{t}_h^S(\gamma)$,thus having $\beta_l < 1$ and $\beta_h = 0$ as the equilibrium rejection probability. While a higher γ in the region of $\gamma > \gamma^{II}$ will induce D_h to propose t_h^S closer to t_h^C, we cannot rule out the possibility of having β_l rise again in response to such an increase in γ, possibly back to $\beta_l = 1$.[45] This is because the benefit from proposing t_h^S closer to t_h^C may outweigh the cost of a higher rejection probability that results from it, inducing D_h to propose $t_h^S = \hat{t}_h^S(\gamma)$. If an improvement in C's signal accuracy reaches another level, denoted by γ^{III}, then it becomes impossible to have the equilibrium value of β_l rise back to $\beta_l = 1$ in response to a further increase in γ. This is because $\hat{t}_h^S(\gamma)$ is reduced to t_h^C when $\gamma = \gamma^{III}$, as shown in Figure 3, which in turn implies that $t_h^S \gg \hat{t}_h^S(\gamma)$, thus $\beta_l < 1$.

When C's signal becomes almost perfect with $\gamma \to 1$, the settlement proposal approaches to the Pareto efficient proposal, t_h^C. Even in such a case, note that $C's$ rejection probability remains strictly positive ($\beta_l > 0$) to deter D_l from imitating D_h's proposal.

5.3 Numerical Example with a Trade Model

We conduct a numerical analysis in this subsection based on a trade model often used in the analysis of trade policy. Consider a two-good (m and x) two-country (D and C) trade model with linear demand and supply.[46] The demand functions of each country are as follows:

$$D_m^D(p_m^D) = 1 - p_m^D,\ D_x^D(p_x^D) = 1 - p_x^D,\ D_m^C(p_m^C) = 1 - p_m^C,\ D_x^C(p_x^C) = 1 - p_x^C, \quad (12)$$

where p_g^i denotes the price of goods g in country i. Specific import tariffs τ_D and τ_C are chosen by D and C, respectively. These are assumed to be the only trade policy instruments. In particular, $p_m^D = p_m^C + \tau_D$ and $p_x^D = p_x^C - \tau_C$.

D and C produce m and x using the following supply functions:

$$Q_m^D(p_m^D) = p_m^D,\ Q_x^D(p_x) = bp_x^D,\ Q_m^C(p_m^C) = bp_m^C,\ Q_x^C(p_x) = p_x^C. \quad (13)$$

Assuming $b > 1$, D becomes a natural importer of m and a natural exporter of x.

[45] We have confirmed that such a non-monotonic movement of β_l in response to an increase in $\gamma \in (\gamma^{II}, \gamma^{III})$ indeed arises through a numerical analysis, as discussed in the following subsection.

[46] Many studies use this simple political-economy model to analyze various issues of trade agreements that arise from private political pressure for protection, including Bagwell and Staiger (2005) and Beshkar (2010a, 2016). The following representation of such a political-economy model directly comes from Beshkar (2016).

Under this specification, the politically-weighted government payoff from the importing sector in D is given by

$$u(\tau_D;\theta_D) = \frac{1}{(3+b)^2}\left\{\frac{1}{2}(1+b)^2 + 2\theta_D + [2\theta_D(1+b)-4]\tau_D \right.$$
$$\left. + \left[\frac{1+\theta_D}{2}(1+b)^2 - 2(3+b)(1+b)\right]\tau_D^2\right\}, \tag{14}$$

where $\theta_D \geq 1$ denotes D's political pressure from the import competing industry. D's payoff from the exporting sector is a function of C's import tariff τ_C:

$$v(\tau_C) = \frac{1}{(3+b)^2}\left\{\frac{(1+b)^2}{2} + 2b + 2(1-b)\tau_C + 2(1+b)\tau_C^2\right\}. \tag{15}$$

For the derivation of equations (14) and (15), see Appendix A of Beshkar (2016).

$C's$ politically-weighted payoff from the importing industry $u(\tau_C, \theta_C)$ and his payoff from the exporting industry $v(\tau_D)$ can be defined symmetrically.

Using u and v constructed above, the payoff function of each government can be defined as follows:

$$W^D(\tau_D, \tau_C;\theta_D) = u(\tau_D;\theta_D) + v(\tau_C),\ W^C(\tau_D, \tau_C;\theta_C) = u(\tau_C;\theta_C) + v(\tau_D). \tag{16}$$

Numerical analysis

The parameter values used for the numerical analysis are as follows: $b = 15$, $l = 1.05$, $h = 1.35$. $\theta_D = h$ if the political pressure is high, and $\theta_D = l$, otherwise. θ_C is fixed at l throughout the analysis. C and D expect the arbitration outcome described in Lemma 1 given that $\gamma^A = 0.76$ and $\rho = 0.8$.

This numerical analysis enables us to compare the equilibrium of the one variable case with that of the two variable one. The top graphs on the left side in Figure 4 show how the equilibrium offer on D_h's import tariff level changes in response to an improvement in $C's$ signal quality. As predicted by Proposition 1 and 5, D_h's import tariff proposal stay constant at τ_h^{max} for the one variable case (denoted by a dotted line named "One Var" in Figure 4) but it stays constant only for the region of $\gamma \leq \gamma^I$, then strictly decreases for the two variable case (denoted by a solid line named "Two Var"). The latter tariff proposal stays strictly higher than the former one, confirming the prediction of $\tau_D^S(h) > \tau_h^{max}$.[47]

[47] Recall that $t_h^S \gg t_h^C \gg t_h^{max}$ as shown in Figure 2.

With regard to the rejection probability, the bottom graphs on the left side in Figure 4 show that it (either β_h or β_l) strictly decreases for the one variable case (denoted by two different dotted lines named "One Var") as the quality of C's information improves. In the same graph, $\beta_h = 0$ and $\beta_l = 1$ in the region of $\gamma \in [\gamma^I, \gamma^{II}]$ for the two variable case (named "Two Var"). While it is not easy tell in Figure 4 due to scaling of the graphs, our numerical analysis shows that β_l increases back to $\beta_l = 1$ even after it drops below one in the region of $\gamma \in [\gamma^{II}, \gamma^{III}]$, confirming the realization of such a possibility discussed for Proposition 7. For all levels of signal accuracy, note that the rejection probability (either β_h or β_l) of the two variable case is strictly lower than the one variable one's. Recall that D_h under the two variable case can signal her type more effectively by proposing an import tariff of C that is strictly greater than an efficient level in conjunction with her own import tariff proposal being greater than what she would have proposed under the one variable case, which makes her proposal less attractive for D_l to imitate. Such a proposal induces the rejection probability under the two variable case to be lower than the one under the one variable case, enabling D_h to obtain an higher expected payoff in the settlement bargaining game.

The top graphs on the right side in Figure 4 demonstrates how the arbitration likelihood (i.e., probability of arbitration) against D_l's potential choice of mimicking D_h's equilibrium tariff offer, denoted by $AL(l)$, changes as C's information improves. For the two variable case, this probability of arbitration stays constant until γ reaches γ^I then monotonically increases in response to a further increase in γ. Such a change in the potential (not observed in the equilibrium) arbitration likelihood is necessary to deter D_l from mimicking D_h's equilibrium tariff offer (t_h^S) that is constant until γ reaches γ^I, then monotonically decreases toward t_h^{max} in response to a further increase in γ. For the one variable case, the same probability of arbitration stays constant as D_h's equilibrium tariff proposal is fixed at τ_h^{max}.

In contrast to this potential arbitration likelihood, $AL(h)$ represents the equilibrium arbitration likelihood (i.e. probability of arbitration) against D_h's equilibrium tariff offer. For the one variable case, $AL(h)$ monotonically decreases in response to an increase in γ, as the theory predicts. For the two variable case, $AL(h)$ also monotonically decreases while we could not theoretically prove such a monotonic change in $AL(h)$. It is also worthwhile to note that $AL(h)$ is lower under the two variable case than under the one variable one, which contributes to a higher expected payoff of D_h under the two variable case shown in the bottoms graphs on the right side in Figure 4.

As the theory predicts, all the informational rent from an improvement in C's signal accrues to D_h as she takes all the surplus from the pre-arbitration settlement bargaining game (through her take-or-leave offer), having the expected payoff of D_l (and C) be constant in response to an increase in γ for both cases. D_l's expected payoff under the two variable case is higher that the one under the one variable case as she can offer a Pareto efficient tariff combination only under the two variable case, holding C's payoff constant at the arbitration equivalent payoff.

6 Conclusion

We consider a game of bilateral bargaining over actions in which (i) the cost of disagreement depends on the private *type* of one of the parties and (ii) the uninformed party receives an imperfect signal about this private type before the actual bargaining. We analyze the role of uncertainty about higher-order beliefs in this bargaining game and establish an anti-public-signal result. That is the efficiency of the settlement bargaining outcome is higher if the signal that the uninformed party receives is private rather than public.

Our anti-public-signal result in pre-arbitration negotiations may be contrasted to the *pro-public-signal* result in arbitration models proposed by Beshkar (2010b) and Park (2011). In these studies, publicizing the private information of the defending party relaxes the self-enforceability condition and reduces the level of inefficient punishment that is required to induce truthfulness. Beshkar (2010b) and Park (2011) find the pro-public-signal result in a repeated-game framework in which enforceability of the agreement is improved by introducing an informative public signal.[48]

Understanding caveats associated with our anti-public-signal result is important. First, the imperfect signal of a defendant's type is "secondary information" in the sense that it does not affect the players' expectation about the court ruling (or more generally, the outcome of any hostile engagement resulting from failing to settle) for a given type of the defendant. If the signal can change the expected court ruling conditional on a specific type of the defendant, thus having its own informational value independent from the defendant's type, then our anti-public-signal result may not hold.[49] Second, we do not explore the possibility that the

[48]The main difference between Beshkar (2010b) and Park (2011) lies in the analysis of no public signal case. The former assumes no private signal, but the latter analyzes the case with an imperfect private signal of potentially deviating actions.

[49]The studies that explore how two-sided private information affects arbitration in the law and

court utilizes such secondary information once it is revealed, possibly affecting the ruling. As reviewed by Spier (2007), the arbitration literature on "evidence" studies how various institutional requirements on evidence affect pre-arbitration settlement and ruling.[50] Given that the nature of imperfect signals that we analyze here is different from those explored by existing studies, extending our paper toward this direction may lead to new results.

In relation with this direction of extension, we conclude our paper by discussing robustness of our results with respect to having a possible pre-dispute information gathering/exchanging stage.[51] First, suppose the defendant had the possibility to provide the hard public evidence of her type, under which conditions would she do it or not, and when the resulting game is the same as what is analyzed in our paper. The high-type defendant would have an incentive to provide such a hard public evidence of her type as it will change the common prior belief of D being a high type (denoted by ρ) in her favor, which in turn raises her expected payoff under arbitration, thus her settlement payoff as well. However, all the results of our analysis will remain qualitatively unchanged as long as such a hard evidence is not strong enough to prove her type.[52] Second, if the signal was somewhat costly to acquire, one may wonder whether the complainant would have incentives to acquire it. Because of our simple bargaining setup of D's proposing a take-it-or-leave-it settlement offer, which induces all the gains from settlement (thereby avoiding costly arbitration) to belong to D, C would not have any incentive to acquire costly signals whether it is private or public. Under other bargaining setups, such as alternating-offer bargaining, in which D and C may share the gains from settlement, then C would have incentives to acquire costly private signals, but continue have no incentive to acquire costly public signals.[53] Analyzing these issues in the context of how to design institutions for optimal evidence generation can be a possible direction for extending our paper.

economics literature assume that both sides of private information have independent informational value on the expected court outcome (Schweizer, 1989; Daughety and Reinganum, 1994).

[50] The institutional requirements include the burden of proof, disclosure and discovery as well as admissibility of settlement negotiations at trial. See Spier (2007) for the review of studies on these issues.

[51] We appreciate the comments from one of referees, which raise the following questions to address robustness of our results.

[52] If the hard evidence can prove the defendant's type, then settlement bargaining becomes a game with complete information.

[53] Regardless of how the signals were acquired, a public signal would lose its informational value in the (Divine) equilibrium as we have shown in our paper.

Appendix

A Proofs

A.1 Proof of Lemma 2

Proof. For any given state of the world, if there exists a point on the line connecting $\tau^A(l)$ and $\tau^A(h)$ that is (weakly) preferred by both parties to the lottery between $\tau^A(l)$ and $\tau^A(h)$, then Lemma 2 holds. To prove that there exists such a point for $\theta = h$, define the following notations:

$\tau^D(a;h)$ and $\tau^C(a)$ are τ that respectively satisfies

$$W^D(\tau;\theta=h) = aW^D\left(\tau^A(h);\theta=h\right) + (1-a)W^D\left(\tau^A(l);\theta=h\right),$$
$$W^C(\tau) = aW^C\left(\tau^A(h)\right) + (1-a)W^C\left(\tau_D^A(l)\right).$$

If $\tau^D(a;h) < \tau^C(a)$, then there exists $\tau \in \left(\tau^D(a;h), \tau^C(a)\right)$ with which both parties prefer playing such an action over the lottery between $\tau^A(l)$ and $\tau^A(h)$ that assigns $\tau^A(h)$ with probability a. Therefore, if $\tau^D(a;h) < \tau^C(a)$ for all $a \in (0,1)$, then we prove (2) for $\theta = h$.

We prove that $\tau^D(a;h) < \tau^C(a)$ for all $a \in (0,1)$ as follows. First, conduct the following monotonic transformation of the payoff functions of D and C on $\tau \in [\tau^A(l), \tau^A(h)]$, denoting the resulting functions by $f_D(\tau)$ and $f_C(\tau)$ respectively:

$$f_D(\tau) \equiv \frac{W^D(\tau;\theta=h) - W^D\left(\tau^A(l);\theta=h\right)}{W^D\left(\tau^A(h);\theta=h\right) - W^D\left(\tau^A(l);\theta=h\right)},$$
$$f_C(\tau) \equiv \frac{W^C(\tau) - W^C\left(\tau^A(l)\right)}{W^C\left(\tau^A(h)\right) - W^C\left(\tau^A(l)\right)}.$$

Note that $f_D(\tau)$ and $f_C(\tau)$ strictly increase from 0 to 1 as τ increases from $\tau^A(l)$ to $\tau^A(h)$. Also note that $\tau^{fD}(a)$ defined by τ that satisfies $f_D(\tau) = a$ is equal to $\tau^D(a;h)$. Similarly, $\tau^{fC}(a)$ defined by τ that satisfies $f_C(\tau) = a$ is equal to $\tau^C(a)$. Given these equalities ($\tau^{fD}(a) = \tau^D(a;h)$ and $\tau^{fC}(a) = \tau^C(a)$), if $f_D(\tau) - f_C(\tau) > 0$ for $\tau \in [\tau^A(l), \tau^A(h)]$, then $\tau^D(a;h) < \tau^C(a)$ for all $a \in (0,1)$.

Thus, it remains to prove $f_D(\tau) - f_C(\tau) > 0$ for $\tau \in \left(\tau^A(l), \tau^A(h)\right)$. Recall that both $W^D(\tau; \theta = h)$ and $W^D(\tau; \theta = h) + W^C(\tau)$ are increasing and concave functions in this range of τ_D whereas $W^C(\tau)$ is a decreasing one in τ. If $f_D^{/}(\tau) - f_C^{/}(\tau) > 0$ at $\tau = \tau^A(l)$, $f_D^{/}(\tau) - f_C^{/}(\tau) = 0$ only once at some $\tau \in \left(\tau^A(l), \tau^A(h)\right)$, and $f_D^{/}(\tau) - f_C^{/}(\tau) < 0$ at $\tau = \tau^A(h)$, then we should have $f_D(\tau) - f_C(\tau) > 0$ for $\tau \in \left(\tau^A(l), \tau^A(h)\right)$ as $f_D\left(\tau^A(l)\right) = f_C\left(\tau^A(l)\right) = 0$ and $f_D\left(\tau^A(h)\right) = f_C\left(\tau^A(h)\right) = 1$.

Now, we prove that $f_D^{/}(\tau) - f_C^{/}(\tau) > 0$ at $\tau = \tau^A(l)$, $f_D^{/}(\tau) - f_C^{/}(\tau) = 0$ only once at some $\tau \in \left(\tau^A(l), \tau^A(h)\right)$, and $f_D^{/}(\tau) - f_C^{/}(\tau) < 0$ at $\tau = \tau^A(h)$ as follows. If one draws the graphs of $\frac{\partial W^D(\tau;\theta=h)}{\partial \tau}$ and $-\frac{\partial W^C(\tau)}{\partial \tau}$ on $\tau \in \left[\tau^A(l), \tau^A(h)\right]$, the former graph is located above the latter one, with the value of the slope of the former one being smaller than that of the latter one (this last inequality results from the concavity of $W^D(\tau; \theta = h) + W^C(\tau)$). The graphs of $f_D^{/}(\tau)$ and $f_C^{/}(\tau)$ are obtained by multiplying a certain positive constant value to $\frac{\partial W^D(\tau;\theta=h)}{\partial \tau}$ and $-\frac{\partial W^C(\tau)}{\partial \tau}$, respectively, so that the resulting areas below $f_D^{/}(\tau)$ and $f_C^{/}(\tau)$ over $\tau \in \left[\tau^A(l), \tau^A(h)\right]$ are both equal to one. Because $W^D(\tau; \theta = h)$ increases more than $W^C(\tau)$ decreases over $[\tau^A(l), \tau^A(h)]$, such a constant value to be multiplied has a smaller value for $\frac{\partial W^D(\tau;\theta=h)}{\partial \tau}$ than the one for $-\frac{\partial W^C(\tau)}{\partial \tau}$. This implies that the value of the slope of $f_D^{/}(\tau)$ is smaller than that of $f_C^{/}(\tau)$, which in turn implies that $f_D^{/}(\tau)$ and $f_C^{/}(\tau)$ cross only once over $[\tau^A(l), \tau^A(h)]$, if they ever cross at all. $f_D^{/}(\tau) - f_C^{/}(\tau) > 0$ at $\tau = \tau^A(l)$ because $f_D^{/}(\tau) - f_C^{/}(\tau) \leq 0$ contradicts with both areas below $f_D^{/}(\tau)$ and $f_C^{/}(\tau)$ being equal to one. Given $f_D^{/}(\tau) - f_C^{/}(\tau) > 0$ at $\tau = \tau^A(l)$, $f_D^{/}(\tau) - f_C^{/}(\tau) \geq 0$ at $\tau = \tau^A(h)$ also lead to a contradiction with both areas below $f_D^{/}(\tau)$ and $f_C^{/}(\tau)$ being equal to one.

One can prove the same result for the case of $\theta = l$ using a similar logic. □

A.2 Proof of Lemma 4

To prove this lemma, we first formally define a Perfect Bayesian Equilibrium (PBE) and a Universally Divine PBE. Then, we characterize both separating and pooling PBEs of this game, with which we prove that a PBE survives the Divinity refinement if and only if it is a fully-separating equilibrium that maximizes D_h's expected payoff among all PBEs.

Definition A1. The strategy profile $(\alpha_l(.), \alpha_h(.), \beta_l(.), \beta_h(.))$ is a PBE *iff*

1. [Incentives of C_j]

If $0 < \beta_j\left(\tau^S\right) < 1$, then C_j is indifferent between settlement at τ^S and litigation, thus satisfying the incentive compatibility condition in (5)

If $\beta_j\left(\tau^S\right) = 1$ $\left(\beta_j\left(\tau^S\right) = 0\right)$, then C_j at least weakly prefers litigation (settlement at τ^S).

2. [Incentives of D_j] For any τ^S for which $0 < \alpha_j\left(\tau^S\right) < 1$, we should have

$$\tau^S = \arg\max_{\tau} \begin{pmatrix} \left[\gamma\left(1-\beta_j(\tau)\right) + (1-\gamma)\left(1-\beta_i(\tau)\right)\right] W^D(\tau; \theta = j) \\ + \left[\gamma\beta_j(\tau) + (1-\gamma)\beta_i(\tau)\right] W_L^D(j) \end{pmatrix},$$

where $i \neq j \in \{l, h\}$.

3. [Consistency of beliefs] If $\alpha_l\left(\tau^S\right) > 0$ or $\alpha_h\left(\tau^S\right) > 0$, then $\Pr\left(\theta = l | \theta^C = l, \tau^S\right)$ and $\Pr\left(\theta = l | \theta^C = h, \tau^S\right)$ are defined as in (5) following the Bayes' rule.

Definition A2. Denote D_h's (D_l's) expected equilibrium payoff by EW^{D_h} (EW^{D_l}). Define $B_h(\tau')$ and $B_l(\tau')$ as a subset of best-response litigation strategies, $(\beta_l(\tau'), \beta_h(\tau'))$ of C on D's off-the-equilibrium settlement proposal, τ', that respectively satisfies the following inequalities:

$$\begin{aligned} &\left[\gamma\beta_h(\tau') + (1-\gamma)\beta_l(\tau')\right] W_L^D(h) + \\ \left\{\gamma\left[1-\beta_h(\tau')\right] + (1-\gamma)\left[1-\beta_l(\tau')\right]\right\} &W^D\left(\tau', h\right) > EW^{D_h}, \\ &\left[(1-\gamma)\beta_h(\tau') + \gamma\beta_l(\tau')\right] W_L^D(l) + \\ \left\{(1-\gamma)\left[1-\beta_h(\tau')\right] + \gamma\left[1-\beta_l(\tau')\right]\right\} &W^D\left(\tau', l\right) > EW^{D_l}. \end{aligned}$$

Similarly, $\bar{B}_h(\tau')$ and $\bar{B}_l(\tau')$ are defined as a subset of best-response strategies of C on D's proposal, τ', that respectively satisfies the above inequalities with weak inequalities. For any off-the-equilibrium settlement proposal, τ', C believes that D_h (D_l) proposed τ' *iff* $B_h(\tau') \supset \bar{B}_l(\tau')$ ($B_l(\tau') \supset \bar{B}_h(\tau')$). A PBE of our dispute settlement game is *Universally Divine iff* it is supported by such belief of C about any off-equilibrium proposal.

Characterization of PBEs:

C1. Under any separating PBE, i.e., a PBE entailing $\exists \tau^S$ with $\alpha_l\left(\tau^S\right) > 0$ and $\alpha_h\left(\tau^S\right) = 0$, $\tau^S = \tau_l^{max}$ and $\beta_l\left(\tau_l^{max}\right) = \beta_h\left(\tau_l^{max}\right) = 0$.

C2. Under any separating PBE, any $\tau^S (\neq \tau_l^{max})$ with $\alpha_h\left(\tau^S\right) > 0$ and a positive settlement probability (i.e., $\beta_l\left(\tau^S\right) < 1$ or $\beta_h\left(\tau^S\right) < 1$) belongs to $\left[\tau_h^{min}, \tau_h^{max}\right]$.

C3. Under any separating PBE, any $\tau^S (\neq \tau_l^{max})$ with $\alpha_h\left(\tau^S\right) > 0$ that belongs to $\left[\tau_h^{min}, \tau_h^{max}\right)$ uniquely defines $\beta_l\left(\tau^S\right)$ and $\beta_h\left(\tau^S\right)$ by (3) holding with equality and Lemma 3.

C4. Among all separating PBEs, the separating PBE that maximizes the expected payoff of D_h is the one with $\alpha_h\left(\tau_b\right) = 1$, $\alpha_l\left(\tau_l^{max}\right) = 1$, and $\beta_l\left(\tau_b\right)$ and $\beta_h\left(\tau_b\right)$ being uniquely determined by (3) holding with equality and Lemma 3 applying even for $\alpha_l\left(\tau_h^{max}\right) = 0$.

C5. Under a pooling PBE, i.e., a PBE with $\alpha_l\left(\tau^S\right) > 0$ and $\alpha_h\left(\tau^S\right) > 0$ for $\forall \tau^S$ that is played with a positive probability along the equilibrium path, $\tau^S < \tau_h^{max}$.

Proving C1) Consider a proposal under a separating PBE, $\tau^S \neq \tau_l^{max}$ with $\alpha_l\left(\tau^S\right) > 0$ and $\alpha_h\left(\tau^S\right) = 0$. Then, if $\tau^S < \tau_l^{max}$, $\beta_l\left(t^S\right) = \beta_h\left(t^S\right) = 0$, and if $\tau^S > \tau_l^{max}$, $\beta_l\left(\tau^S\right) = \beta_h\left(\tau^S\right) = 1$. Note that the expected payoff of D_l from proposing such $\tau^S \neq \tau_l^{max}$ is strictly smaller that its payoff from proposing $\tau^S = \tau_l^{max}$ with $\beta_l\left(\tau_l^{max}\right) = \beta_h\left(\tau_l^{max}\right) = 0$ because $W^D\left(\tau_l^{max}; l\right) > W_A^D\left(l\right)$ and $W^D\left(\tau_l^{max}; l\right) > W^D\left(\tau^S; l\right)$ for $\tau^S < \tau_l^{max}$. Now it remains to show that $\tau^S = \tau_l^{max}$ with $\alpha_l\left(\tau_l^{max}\right) > 0$, $\alpha_h\left(\tau_l^{max}\right) = 0$, and $\beta_l\left(\tau_l^{max}\right) > 0$ or $\beta_h\left(\tau_l^{max}\right) > 0$ cannot be a part of a separating PBE. To show this by contradiction, first assume that it is a part of a separating equilibrium. Then, there exists $\tau' < \tau_l^{max}$ that is sufficiently close to τ_l^{max} such that the expected payoff of D_l from proposing τ' with $\beta_l\left(\tau'\right) = \beta_h\left(\tau'\right) = 0$ is strictly greater than the expected payoff of D_l from proposing $\tau^S = \tau_l^{max}$ with $\beta_l\left(\tau_l^{max}\right) > 0$ or $\beta_h\left(\tau_l^{max}\right) > 0$. This is because $W^D\left(\tau_l^{max}; l\right) > W_A^D\left(l\right)$, thus leading to contradiction.

Proving C2) If τ^S does not belong to $\left[\tau_h^{min}, \tau_h^{max}\right]$, then either D_h or C will strictly prefer litigating over settlement with τ^S. D_h strictly prefers arbitration over settlement with $\tau^S < \tau_h^{min}$ that entails a positive probability of settlement, thus proposing such $\tau^S < \tau_h^{min}$ is strictly dominated by a proposal that will surely invoke arbitration, such as $\tau^S > \tau_h^{max}$. C strictly prefers arbitration over settlement with $\tau^S > \tau_h^{max}$, then such τ^S will entail zero probability of settlement.

Proving C3) Under a separating PBE, consider $\tau^S (\neq \tau_l^{max})$ with $\alpha_h\left(t^S\right) > 0$ that belong to $\left[\tau_h^{min}, \tau_h^{max}\right)$. For such τ^S, first note that $\beta_l\left(\tau^S\right)$ and $\beta_h\left(\tau^S\right)$ (with $\beta_l\left(\tau^S\right) > 0$) need to satisfy D_l's incentive constraint in (3) with equality: if D_l's incentive constraint in (3) holds with a strictly inequality, then $\alpha_l\left(\tau^S\right)=0$, implying that C, who correctly believes that only D_h would propose such τ^S, will settle for sure with $W^C\left(\tau^S\right)>W_A^C(h)$, which in turn contradicts with D_l's incentive constraint in (3) holding with a strictly inequality. The same logic also implies that $\alpha_l\left(t^S\right)>0$ for $\tau^S(\neq\tau_l^{max})\in\left[\tau_h^{min},\tau_h^{max}\right)$ with $\alpha_h\left(t^S\right)>0$ under a separating PBE. Because $W^D\left(\tau^S;l\right)>W^D\left(\tau_l^{max};l\right)>W_A^D(l)$, (3) uniquely defines $\gamma\beta_l\left(\tau^S\right)+(1-\gamma)\beta_h\left(\tau^S\right)$ to be a rational value $\in(0,1)$. With Lemma 3 and $\gamma\in(0.5,1)$, such a requirement uniquely defines $\beta_l\left(\tau^S\right)$ and $\beta_h\left(\tau^S\right)$.

Proving C4) With regard to the choice among all separating PBEs, $D_h's$ constrained maximization in (4) involves choosing $\tau^S\in\left[\tau_h^{min},\tau_h^{max}\right]$ with (3) and Lemma 3 together uniquely defining a pair of $\beta_l\left(\tau^S\right)$ and $\beta_h\left(\tau^S\right)$: while Lemma 3 may not apply to $\tau^S=\tau_h^{max}$ with $\alpha_l\left(\tau_h^{max}\right)=0$, we can show that the uniquely defined pair of $\beta_l\left(\tau_h^{max}\right)$ and $\beta_h\left(\tau_h^{max}\right)$ with Lemma 3 holding is the one that results from $D_h's$ constrained maximization in (4) when $\tau_b=\tau_h^{max}$. Because $W^D\left(\tau^S;l\right)>W^D\left(\tau_l^{max};l\right)>W_A^D(l)$ for $\tau^S\in\left[\tau_h^{min},\tau_h^{max}\right]$ and $\partial W^D\left(\tau^S;l\right)/\partial\tau^S>0$, an increase in τ^S requires a corresponding increase in $\gamma\beta_l(\tau^S)+(1-\gamma)\beta_h(\tau^S)$, which raises the cost associated with increasing τ^S by raising the likelihood of costly arbitration being invoked. Since $\beta_l(\tau)\in[0,1)$ implies $\beta_h(\tau)=0$, and $\beta_h(\tau)\in(0,1]$ implies $\beta_l(\tau)=1$, as discussed in Lemma 3, this means that either only an increase in $\beta_l(\tau^S)$ or only an increase in $\beta_h(\tau^S)$. If there is a unique value of τ_b, then a separating PBE that maximizes the expected payoff of D_h should entail $\alpha_h(\tau_b)=1$. Even if there exist multiple values of τ_b, a separating PBE with $\alpha_h(\tau_b)=1$ on one of these multiple values is one of such PBEs that maximizes the expected payoff of D_h.

While the above paragraph completes the proof for C4, we can provide the following characterization of C's equilibrium strategy. If $\tau_b=\tau_h^{max}$, then $\alpha_l\left(\tau_h^{max}\right)=0$ because $\alpha_l(\tau_h^{max})>0$ will induce $\beta_l\left(\tau^S\right)=\beta_h\left(\tau^S\right)=1$. If $\tau_b<\tau_h^{max}$, $\alpha_l(\tau_b)\left(=1-\alpha_l\left(\tau_l^{max}\right)\right)>0$ is defined by C's incentive compatibility condition in (5). First note that $W_A^C(l)>W^C\left(\tau^S\right)\geq W_A^C(h)$ for $\tau^S\in\left[\tau_h^{min},\tau_h^{max}\right]$ and $\Pr\left(\theta=l|\theta^C=j,\tau^S\right)=0$ for $\alpha_l\left(\tau^S\right)=0$ with $\partial\Pr\left(\theta=l|\theta^C=j,\tau^S\right)/\partial\alpha_l\left(\tau^S\right)>0$. For $\tau_b<\tau_h^{max}$, then there exists a unique value of $\alpha_l(\tau_b)$ (>0), with which the incentive compatibility condition in (5) is satisfied for C_{θ^C} with $0<\beta_{\theta^C}(\tau_b)<1$. Such a strategy profile constitutes a separating PBE that maximizes D_h's expected payoff. If $\beta_h(\tau_b)=0$ and $\beta_l(\tau_b)=1$, we can have $\alpha_l(\tau_b)$ be defined by the incentive compatibility condition in (5) for C_h without loss of generality.

Proving C5) Consider a pooling PBE with $\exists\tau^S\geq\tau_h^{max}$ with $\alpha_l\left(\tau^S\right)>0$ and $\alpha_h\left(\tau^S\right)>0$. Once such τ^S is proposed, then C will litigate it with probability one because $W^C\left(\tau^S\right)\left(\leq W_A^C(h)\right)$ is strictly smaller than C's expected payoff from arbitration. Given that D_l's expected payoff from playing any pooling PBE is strictly greater than $W_A^D(l)$, which is shown below to hold, $\alpha_l\left(\tau^S\right) > 0$ with $\tau^S \geq \tau_h^{max}$ cannot be a part of the pooling PBE because it will violate the incentive constraint of D_l that needs to be satisfied for any PBE.

To show that D_l's expected payoff from playing any pooling PBE is strictly greater than $W_A^D(l)$, we provide further characterization of a pooling PBE as follows. First consider a pure-strategy pooling equilibrium with $\alpha_h(\tau^S) = \alpha_l(\tau^S) = 1$. Let $\rho_{\theta^C}^*$ denote C's updated belief of the likelihood of D being a high type after observing θ^C. Moreover, let $\tau_{pool}\left(\rho_{\theta^C}^*\right)$ denote a proposal that makes C_{θ^C} indifferent between settlement and arbitration when all types of D pool, namely,

$$\left(1-\rho_{\theta^C}^*\right) W_A^C(l) + \rho_{\theta^C}^* W_A^C(h) \equiv W^C\left(\tau_{pool}\left(\rho_{\theta^C}^*\right)\right),$$

with

$$\rho_l^* = \Pr\left(\theta = h | \theta^C = l, \tau^S\right) = \frac{(1-\gamma)\rho}{\gamma(1-\rho)+(1-\gamma)\rho},$$
$$\rho_h^* = \Pr\left(\theta = h | \theta^C = h, \tau^S\right) = \frac{\gamma\rho}{(1-\gamma)(1-\rho)+\gamma\rho},$$

according to Bayes' rule.

As γ increases, $\tau_{pool}\left(\rho_h^*\right)$ increases monotonically toward $\tau_h^{\max}$ and $\tau_{pool}\left(\rho_l^*\right)$ decreases monotonically toward $\tau_l^{\max}$. With regard to C's strategy, a pooling equilibrium proposal of $\tau^S \in \left(\tau_{pool}\left(\rho_l^*\right), \tau_{pool}\left(\rho_h^*\right)\right)$ entails $\beta_l\left(\tau^S\right) = 1$ and $\beta_h\left(\tau^S\right) = 0$; $\tau^S < \tau_{pool}\left(\rho_l^*\right)$ entails $\beta_l\left(\tau^S\right) = 0$ and $\beta_h\left(\tau^S\right) = 0$; $\tau^S > \tau_{pool}\left(\rho_h^*\right)$ entails $\beta_l\left(\tau^S\right) = 1$ and $\beta_h\left(\tau^S\right) = 1$; $\tau^S = \tau_{pool}\left(\rho_l^*\right)$ entails $\beta_l\left(\tau^S\right) \in [0,1]$ and $\beta_h\left(\tau^S\right) = 0$; and $\tau^S = \tau_{pool}\left(\rho_h^*\right)$ entails $\beta_l\left(\tau^S\right) = 1$ and $\beta_h\left(\tau^S\right) \in [0,1)$. Because D_l and D_h can attain $W^D\left(\tau_l^{min}; l\right)$ and $W^D\left(\tau_h^{min}; h\right)$, respectively by proposing an action that will surely invoke arbitration, a pooling equilibrium proposal must have $\tau^S \geq \tau_h^{min}$. This implies that a pooling equilibrium proposal $\tau^S \in \left[\tau_h^{min}, \tau_{pool}\left(\rho_h^*\right)\right]$, which in turn requires γ to be sufficiently high to have $\tau_{pool}\left(\rho_h^*\right) > \tau_h^{min}$: recall that $\tau_{pool}(\rho) < \tau_h^{\min}$ from Corollary 1. Given this condition on γ is satisfied, then a pooling equilibrium must entail $\beta_l\left(\tau^S\right) = 1$ and $\beta_h\left(\tau^S\right) \in [0,1)$. Because $\tau^S \geq \tau_h^{min}$ with $W^D\left(\tau^S; l\right) > W_A^D(l)$, $\beta_l\left(\tau^S\right) = 1$ and $\beta_h\left(\tau^S\right) \in [0,1)$, D_l's expected payoff from playing any pooling PBE is strictly greater than $W_A^D(l)$.

Now consider the possibility of having a mixed-strategy pooling equilibrium with a multiple proposal, τ_i^S, with $\alpha_l\left(\tau_i^S\right) > 0$ and $\alpha_h\left(\tau_i^S\right) > 0$. First note that $\tau_i^S \geq \tau_h^{min}$ because D_h would strictly prefer arbitration, otherwise. Also, note that the expected payoff from playing any τ_i^S should yield the same expected payoff for D of any given type. This requires assigning a strictly higher arbitration probability for a higher value τ_i^S.

For a given pair of $\alpha_l\left(\tau_i^S\right) > 0$ and $\alpha_h\left(\tau_i^S\right) > 0$, note that both $\tau_{pool}\left(\rho_h^*\right)$ and $\tau_{pool}\left(\rho_l^*\right)$ increase (decrease) monotonically toward $\tau_h^{\max}$ ($\tau_l^{\max}$) with $\tau_{pool}\left(\rho_h^*\right) > \tau_{pool}\left(\rho_l^*\right)$ as $\alpha_h\left(\tau_i^S\right)/\alpha_l\left(\tau_i^S\right)$ increases (decreases). Also note that a pooling equilibrium proposal of $\tau_i^S \in \left(\tau_{pool}\left(\rho_l^*\right), \tau_{pool}\left(\rho_h^*\right)\right)$ entails $\beta_l\left(\tau_i^S\right) = 1$ and $\beta_h\left(\tau_i^S\right) = 0$; $\tau_i^S < \tau_{pool}\left(\rho_l^*\right)$ entails $\beta_l\left(\tau_i^S\right) = 0$ and $\beta_h\left(\tau_i^S\right) = 0$; $\tau_i^S > \tau_{pool}\left(\rho_h^*\right)$ entails $\beta_l\left(\tau_i^S\right) = 1$ and $\beta_h\left(\tau_i^S\right)$; $\tau_i^S = \tau_{pool}\left(\rho_l^*\right)$ entails $\beta_l\left(\tau_i^S\right) \in [0,1]$ and $\beta_h\left(\tau_i^S\right) = 0$; and $\tau_i^S = \tau_{pool}\left(\rho_h^*\right)$ entails $\beta_l\left(\tau_i^S\right) = 1$ and $\beta_h\left(\tau_i^S\right) \in [0,1)$. Denote the lowest value among τ_i^S by τ_{min}^S. Then, $\alpha_h\left(\tau_{min}^S\right)/\alpha_l\left(\tau_{min}^S\right)$ should take the lowest value among $\alpha_h\left(\tau_i^S\right)/\alpha_l\left(\tau_i^S\right)$ to assign a strictly higher arbitration probability for a higher value τ_i^S. This is because $\tau_{min}^S \geq \tau_h^{min}$ and $\alpha_h\left(\tau_{min}^S\right)/\alpha_l\left(\tau_{min}^S\right) < \rho$, which in turn implies $\tau_{min}^S \in \left(\tau_{pool}\left(\rho_l^*\right), \tau_{pool}\left(\rho_h^*\right)\right]$ entailing $\beta_l\left(\tau_i^S\right) = 1$ and $\beta_h\left(\tau_i^S\right) = 0$. To assign a strictly higher arbitration probability for a higher value τ_i^S, then $\tau_i^S = \tau_{pool}\left(\rho_h^*\right)$ entails $\beta_l\left(\tau_i^S\right) = 1$ and $\beta_h\left(\tau_i^S\right) \in [0,1)$ for any $\tau_i^S > \tau_{min}^S$. Thus, the highest value among τ_i^S, denoted by τ_{max}^S, should be equal to $\tau_{pool}\left(\rho_h^*\right)$ with $\alpha_h\left(\tau_{max}^S\right)/\alpha_l\left(\tau_{max}^S\right)$ being the highest value among $\alpha_h\left(\tau_i^S\right)/\alpha_l\left(\tau_i^S\right)$. As the expected payoff from playing any τ_i^S should yield the same expected payoff for D of any given type, we can focus on the expected payoff of D_l from playing τ_{max}^S ($> \tau_h^{min}$) entailing $\beta_l\left(\tau_{max}^S\right) = 1$ and $\beta_h\left(\tau_{max}^S\right) \in [0,1)$. Playing such a pooling equilibrium yields D_l the expected payoff that is strictly greater than $W_A^D(l)$ because $W^D\left(\tau_{max}^S; l\right) > W_A^D(l)$ with $\tau_{max}^S > \tau_h^{min}$.

Now, we prove Lemma 4 by proving i) the only separating PBE that satisfies the Universal Divinity criterion is the separating equilibrium that maximizes the expected payoff of D_h (that is characterized in C4), **and by proving ii)** no pooling PBE satisfies the Universal Divinity criterion.

First, consider a separating PBE that is not maximizing D_h's expected payoff. The expected payoff of D_l under such a PBE (or any separating PBE), denoted by EW^{D_l}, is equal to $W^D\left(\tau_l^{max}; l\right)$: if $EW^{D_l} > W^D\left(\tau_l^{max}; l\right)$, D_l has a strict incentive to assign $\alpha_l\left(\tau_l^{max}\right) = 0$, thus invalidating the separating behavior characterized above in C1; if $EW^{D_l} < W^D\left(\tau_l^{max}; l\right)$, D_l has a strict incentive to assign

$\alpha_l\left(\tau_l^{max}\right) = 1$, thus having $EW^{D_l} = W^D\left(t_l^{max}; l\right)$. Denote the expected payoff of D_h from playing such a separating PBE by $EW^{D_h}(< EW^{D_h}(\tau_b))$, where $EW^{D_h}(\tau_b)$ represents the maximized expected payoff of D_h. Now consider a possible off-the-equilibrium deviation in which $\tau^S = \tau^{/} > \tau_b$ (or $\tau^S = \tau^{/} < \tau_b$ if $\tau_b = \tau_h^{max}$) with $\tau^{/} \in \left[\tau_h^{min}, \tau_h^{max}\right)$. In particular, $\tau^{/}$ is close enough to τ_b so that the expected payoff of D_h from playing $\alpha_h\left(\tau^{/}\right) = 1$ with $\beta_l\left(\tau^{/}\right)$ and $\beta_h\left(\tau^{/}\right)$ being respectively equal to $\beta_l^*\left(\tau^{/}\right)$ and $\beta_h^*\left(\tau^{/}\right)$ that are uniquely determined by (3) holding with equality and Lemma 3, which we denote by $EW^{D_h}\left(\tau^{/}\right)$, satisfies the following inequalities: $EW^{D_h} < EW^{D_h}\left(\tau^{/}\right) < EW^{D_h}(\tau_b)$.

Note that $\beta_l\left(\tau^{/}\right) \in [0,1)$ implies $\beta_h\left(\tau^{/}\right) = 0$, and $\beta_h\left(\tau^{/}\right) \in (0,1]$ implies $\beta_l\left(\tau^{/}\right) = 1$ for $\tau^{/} \in \left[\tau_h^{min}, \tau_h^{max}\right)$ from C3. This enables us to represent the strategy of C by a single variable, $\beta_S(\tau^{/}) \equiv \beta_l(\tau^{/}) + \beta_h(\tau^{/}) \in [0,2]$. In a similar manner, define $\beta_S^*\left(\tau^{/}\right) \equiv \beta_l^*\left(\tau^{/}\right) + \beta_h^*\left(\tau^{/}\right)$. Also, define $\hat{\beta}_S\left(\tau^{/}\right) \equiv \hat{\beta}_l\left(\tau^{/}\right) + \hat{\beta}_h\left(\tau^{/}\right)$ as the value of $\beta_S\left(\tau^{/}\right)$ that makes the expected payoff of D_h from playing $\alpha_h\left(\tau^{/}\right) = 1$ with $\hat{\beta}_S\left(\tau^{/}\right)$ be equal to EW^{D_h}. Thus, $B_h\left(\tau^{/}\right) = \{\beta_S\left(\tau^{/}\right) < \hat{\beta}_S\left(\tau^{/}\right)\}$. $\bar{B}_l\left(\tau^{/}\right) = \{\beta_S\left(\tau^{/}\right) \leq \beta_S^*\left(\tau^{/}\right)\}$ because the expected payoff of D_l from playing $\alpha_l\left(\tau^{/}\right) = 1$ with $\beta_S = \beta_S^*(\tau^{/})$, denoted by $EW^{D_l}\left(\tau^{/}\right)$, is equal to $W^D\left(\tau_l^{\max}; l\right) = EW^{D_l}$. Because $\hat{\beta}_S\left(\tau^{/}\right) > \beta_S^*(\tau^{/})$, $B_h\left(\tau^{/}\right) \supset \bar{B}_l\left(\tau^{/}\right)$, inducing C to believe that D_h has deviated to proposing $\tau^{/} \in \left[\tau_h^{min}, \tau_h^{max}\right)$. Because $W^C(\tau^{/}) > W_A^C(h)$, C will accept such a deviation proposal $\tau^{/}$, creating a deviation incentive for D_h. This proves that a separating PBE that is not maximizing D_h's expected payoff is not a Divine equilibrium.

For the first part, it remains to show that the separating equilibrium that maximizes the expected payoff of D_h characterized in C4, does satisfy the Divinity criterion against any off-the-equilibrium action of $\tau^S = \tau^{/}(\neq \tau_b) \in \left[\tau_h^{min}, \tau_h^{max}\right)$. Even though it is possible to have $\tau^{/} = \tau_h^{max}$ if $\tau_b < \tau_h^{max}$, we can focus on $\tau^{/} < \tau_h^{max}$ because $\tau_b = \tau_h^{max}$ as shown in Proposition 1. With respect to an off-the-equilibrium action of $\tau^S = \tau^{/}(\neq \tau_b) \in \left[\tau_h^{min}, \tau_h^{max}\right)$, define $\beta_S^*\left(\tau^{/}\right)$ be the value of $\beta_S\left(\tau^{/}\right)$ that makes the expected payoff of D_l from $\tau^S = \tau^{/}$ be equal to $W^D\left(\tau_l^{max};l\right)$. Note that $\beta_S^*\left(\tau^/\right)$ is uniquely determined by (3) holding with equality for $\tau^S=\tau^/$. Now, define $\hat{\beta}_S\left(\tau^/\right)$ as the value of $\beta_S\left(\tau^/\right)$ that makes the expected payoff of D_h from playing $\alpha_h\left(\tau^/\right)=1$ with $\beta_S\left(\tau^/\right)=\hat{\beta}_S\left(\tau^/\right)$ be equal to $EW^{D_h}\left(\tau_b\right)$. Note that $\hat{\beta}_S\left(\tau^/\right)<\beta_S^*\left(\tau^/\right)$ because the expected payoff of D_h from playing $\alpha_h\left(\tau^/\right)=1$ with $\beta_S\left(\tau^/\right)=\beta_S^*\left(\tau^/\right)$ is smaller that $EW^{D_h}\left(\tau_b\right)$. This implies $B_l\left(\tau^/\right)=\{\beta_S\left(\tau^/\right)<\beta_S^*\left(\tau^/\right)\}\supset\bar{B}_h\left(\tau^/\right)=\{\beta_S\left(\tau^/\right)\leq\hat{\beta}_S\left(\tau^/\right)\}$, which in turn induces C to believe that D_l (not D_h) deviates by proposing $\tau^/$. Thus, C will litigate against $\tau^/$ with $\beta_S\left(\tau^/\right)=2$, eliminating any incentive for D to deviate from the separating equilibrium that maximizes the expected payoff of D_h.

For the second part, we prove that there exists no pooling PBE that satisfies the Universal Divinity criterion. Given C5 above, we can consider a deviation from a pooling equilibrium, which can be either a pure-strategy pooling proposal, τ^S with $\alpha_l\left(\tau^S\right)=\alpha_h\left(\tau^S\right)=1$ or a mixed-strategy pooling proposal with $\tau^S=\tau^S_{max}$ ($>\tau_h^{min}$). With regard to a deviation from a mixed strategy pooling equilibrium, we can focus on a deviation from $\tau^S=\tau^S_{max}$ because D_j's expected payoff from offering any mixed strategy pooling proposal, τ_i^S with $\alpha_j\left(\tau_i^S\right)>0$, is identical to D_j's expected payoff from the mixed strategy pooling equilibrium for $j=l$ or h. With regard to such a proposal τ^S, recall that a pooling equilibrium must entail $\beta_l\left(\tau^S\right)=1$ and $\beta_h\left(\tau^S\right)\in[0,1)$ from the proof for C5. Denote such τ^S by τ_{pool}, and denote the expected payoffs of D_l and D_h of a pooling PBE of proposing τ_{pool} by $EW^{D_l}_{pool}$ and $EW^{D_h}_{pool}$, respectively. Now consider a possible off-the-equilibrium deviation $\tau^/$ that belongs to $\left(\tau_{pool},\tau_h^{max}\right)$, which is a non-empty set by C5. Note that the expected payoff of D_h (D_l) from playing $\alpha_h\left(\tau^/\right)=1$ ($\alpha_l\left(\tau^/\right)=1$) with $\beta_S\left(\tau^/\right)=\beta_S^*\left(\tau_{pool}\right)$, the probability of arbitration associated with τ_{pool} in the pooling equilibrium, denoted by $EW^{D_h}_{pool}\left(\tau^/\right)$ ($EW^{D_l}_{pool}\left(\tau^/\right)$), satisfies the following inequality: $EW^{D_h}_{pool}<EW^{D_h}_{pool}\left(\tau^/\right)$ ($EW^{D_l}_{pool}<EW^{D_l}_{pool}\left(\tau^/\right)$). Finally, define $\hat{\beta}_S^{D_h}\left(\tau^/\right)$ ($\hat{\beta}_S^{D_l}\left(\tau^/\right)$) as the value of $\beta_S\left(\tau^/\right)$ that makes the expected payoff of D_h (D_l) from playing $\alpha_h\left(\tau^/\right)=1$ ($\alpha_l\left(\tau^/\right)=1$) with $\beta_S\left(\tau^/\right)=\hat{\beta}_S^{D_h}\left(\tau^/\right)$ ($=\hat{\beta}_S^{D_l}\left(\tau^/\right)$)

be equal to $EW^{D_h}_{pool}$ ($EW^{D_h}_{pool}$). Note that $B_h\left(\tau^{/}\right) = \{\beta_S\left(\tau^{/}\right) < \hat{\beta}^{D_h}_S\left(\tau^{/}\right)\}$ and $\bar{B}_l\left(\tau^{/}\right) = \{\beta_S\left(\tau^{/}\right) \leq \hat{\beta}^{D_l}_S\left(\tau^{/}\right)\}$. As shown below, the inequality condition in (6) implies that there exists $\tau^{/} \in (\tau_{pool}, \tau^{max}_h)$ with $\hat{\beta}^{D_h}_S\left(\tau^{/}\right) \equiv \hat{\beta}^{D_h}_h\left(\tau^{/}\right) + 1 > \hat{\beta}^{D_l}_S\left(\tau^{/}\right) \equiv \hat{\beta}^{D_l}_h\left(\tau^{/}\right) + 1$, thus, $B_h\left(\tau^{/}\right) \supset \bar{B}_l\left(\tau^{/}\right)$. Then, C would believe that D_h (not D_l) has deviated to proposing $\tau^{/}$. Because $W^C(\tau^{/}) > W^C_A(h)$, C will accept such a deviation proposal $\tau^{/}$, creating a deviation incentive for D. This proves that any pooling PBE is not a Divine equilibrium.

It remains to prove that the inequality condition in (6) implies that there exists $\tau^{/} \in (\tau_{pool}, \tau^{max}_h)$ with $\hat{\beta}^{D_h}_S\left(\tau^{/}\right) > \hat{\beta}^{D_l}_S\left(\tau^{/}\right)$. With regard to a pooling PBE behavior of proposing τ_{pool}, recall that a pooling equilibrium must entail $\beta_l\left(\tau_{pool}\right) = 1$ and $\beta_h\left(\tau_{pool}\right) \in [0, 1)$ with

$$
\begin{aligned}
EW^{D_h} &= \left[\gamma\beta_h\left(\tau_{pool}\right) + (1-\gamma)\right] W^D_L(h) + \gamma\left[1 - \beta_h\left(\tau_{pool}\right)\right] W^D\left(\tau^S, h\right), \\
EW^{D_l} &= \left[(1-\gamma)\beta_h\left(\tau_{pool}\right) + \gamma\right] W^D_L(l) + (1-\gamma)\left[1 - \beta_h\left(\tau_{pool}\right)\right] W^D\left(\tau_{pool}, l\right).
\end{aligned}
$$

If we can prove that there exists $\tau^{/} \in (\tau_{pool}, \tau^{max}_h)$ having $\hat{\beta}^{D_h}_S\left(\tau^{/}\right) \equiv \hat{\beta}^{D_h}_h\left(\tau^{/}\right) + 1$ and $\hat{\beta}^{D_l}_S\left(\tau^{/}\right) \equiv \hat{\beta}^{D_l}_h\left(\tau^{/}\right) + 1$ be uniquely defined by

$$
\begin{aligned}
EW^{D_h} &= \left[\gamma\hat{\beta}^{D_h}_h\left(\tau^{/}\right) + (1-\gamma)\right] W^D_L(h) + \gamma\left[1 - \hat{\beta}^{D_h}_S\left(\tau^{/}\right)\right] W^D\left(\tau^{/}, h\right), \\
EW^{D_l} &= \left[(1-\gamma)\hat{\beta}^{D_l}_h\left(\tau^{/}\right) + \gamma\right] W^D_L(l) + (1-\gamma)\left[1 - \hat{\beta}^{D_l}_h\left(\tau^{/}\right)\right] W^D\left(\tau^{/}, l\right),
\end{aligned}
$$

respectively, that results in $\hat{\beta}^{D_h}_h\left(\tau^{/}\right) > \hat{\beta}^{D_l}_l\left(\tau^{/}\right)$, then it completes the proof. From the above four equalities (from the first and the third ones, and from the second and the fourth ones, more precisely), we can show that

$$
\begin{aligned}
\hat{\beta}^{D_h}_h\left(\tau^{/}\right) &= \frac{\beta_h\left(\tau_{pool}\right)\left[W^D\left(\tau_{pool}, h\right) - W^D_L(h)\right] + \left[W^D\left(\tau^{/}, h\right) - W^D\left(\tau_{pool}, h\right)\right]}{W^D\left(\tau^{/}, h\right) - W^D_L(h)}, \\
\hat{\beta}^{D_l}_h\left(\tau^{/}\right) &= \frac{\beta_h\left(\tau_{pool}\right)\left[W^D\left(\tau_{pool}, l\right) - W^D_L(l)\right] + \left[W^D\left(\tau^{/}, l\right) - W^D\left(\tau_{pool}, l\right)\right]}{W^D\left(\tau^{/}, l\right) - W^D_L(l)}.
\end{aligned}
$$

By differentiating $\hat{\beta}_h^{D_h}\left(\tau^/\right)$ and $\hat{\beta}_l^{D_l}\left(\tau^/\right)$ with respect to $\tau^/$, we obtain

$$\frac{\partial\hat{\beta}_h\left(\tau^/\right)}{\partial\tau^/} = \frac{\frac{\partial W^D(\tau^/;h)}{\partial\tau^/}\left[1-\beta_h\left(\tau_{pool}\right)\right]}{W^D\left(\tau^/,h\right)-W_L^D\left(h\right)},$$

$$\frac{\partial\bar{\beta}_h\left(\tau^/\right)}{\partial\tau^/} = \frac{\frac{\partial W^D(\tau^/;l)}{\partial\tau^/}\left[1-\beta_h\left(\tau_{pool}\right)\right]}{W^D\left(\tau^/,l\right)-W_L^D\left(l\right)}$$

at $\tau^/ = \tau_{pool}$. If the inequality condition in (6) holds, then $\partial\hat{\beta}_h\left(\tau^/\right)/\partial\tau^/ > \partial\hat{\beta}_l\left(\tau^/\right)/\partial\tau^/$ at at $\tau^/ = \tau_{pool}$, which in turn implies that there exists $\tau^/ \in \left(\tau_{pool}, \tau_h^{max}\right)$ with $\hat{\beta}_S^{D_h}\left(\tau^/\right) > \hat{\beta}_S^{D_l}\left(\tau^/\right)$.

A.3 Proof of Proposition 1

According to Lemma 4, we focus on a separating PBE that maximizes the expected payoff of D_h. In showing that Then, C3 and C4 in the proof for Lemma 4 enable us to focus on the solution to (4), namely τ_b, as the equilibrium proposal of D_h, having (3) holding with equality and $\beta_h\left(1-\beta_l\right)=0$ with $\beta_l > \beta_h$ uniquely define $\beta_l\left(\tau_b\right)$ and $\beta_h\left(\tau_b\right)$: because of the incentive compatibility condition of C that yields Lemma 3, $\beta_l > \beta_h$. According to C2 in the proof for Lemma 4, also note that $\tau_b \in \left[\tau_h^{min}, \tau_h^{max}\right]$.

Then, the first order condition associated with (4) is

$$\begin{aligned} &-\left[\gamma\frac{\partial\beta_h(\tau)}{\partial\tau}+(1-\gamma)\frac{\partial\beta_l(\tau)}{\partial\tau}\right]\left[W^D\left(\tau;h\right)-W_A^D\left(h\right)\right] \quad + \\ &\left\{\gamma\left[1-\beta_h(\tau)\right]+(1-\gamma)\left[1-\beta_l(\tau)\right]\right\}\frac{\partial W^D(\tau;h)}{\partial\tau} \qquad \geq 0, \end{aligned}$$

where

$$\begin{aligned} \beta_l(\tau) &= \frac{W^D(\tau;l)-W^D\left(\tau_l^{\max};l\right)}{\gamma\left[W^D(\tau;l)-W_A^D(l)\right]}, \\ \frac{\partial\beta_l(\tau)}{\partial\tau} &= \frac{\frac{\partial W^D(\tau;l)}{\partial\tau}\left[W^D\left(\tau_l^{\max};l\right)-W_A^D(l)\right]}{\gamma\left[W^D(\tau;l)-W_A^D(l)\right]^2}, \\ \beta_h(t) &= -\frac{\gamma}{(1-\gamma)}+\frac{W^D(\tau;l)-W^D\left(\tau_l^{\max};l\right)}{(1-\gamma)\left[W^D(\tau;l)-W_A^D(l)\right]}, \\ \frac{\partial\beta_h(\tau)}{\partial\tau} &= \frac{\frac{\partial W^D(\tau;l)}{\partial\tau}\left[W^D\left(\tau_l^{\max};l\right)-W_A^D(l)\right]}{(1-\gamma)\left[W^D(\beta_l\tau;l)-W_A^D(l)\right]^2}, \end{aligned}$$

with $\beta_h(1-\beta_l) = 0$; $\partial\beta_l(\tau)/\partial\tau = 0$ if $\beta_l = 1$; and $\partial\beta_h(\tau)/\partial\tau = 0$ if $\beta_h = 0$. If $\beta_h > 0$ and $\beta_l = 1$, then the first order condition can be rewritten into:

$$\begin{array}{ll} -\frac{\partial W^D(\tau;l)}{\partial\tau}\left[W^D(\tau;h) - W^D_A(h)\right] & + \\ \left[W^D(\tau;l) - W^D_A(l)\right]\frac{\partial W^D(\tau;h)}{\partial\tau} & \geq 0 \end{array},$$

which holds with a strict inequality for $\tau \in \left[\tau_h^{min}, \tau_h^{max}\right]$ if the inequality condition in (6) holds. If $\beta_h = 0$ and $\beta_l \leq 1$, then the first order condition can be rewritten into:

$$-\frac{\partial W^D(t;l)}{\partial t}\left[W^D\left(t_l^{\max};l\right) - W^D_A(l)\right]\left[W^D(t;h) - W^D_A(h)\right] + \left\{\frac{\gamma}{(1-\gamma)}\left[W^D(t;l) - W^D_A(l)\right] - \left[W^D(t;l) - W^D\left(t_l^{\max};l\right)\right]\right\}\left[W^D(t;l) - W^D_A(l)\right]\frac{\partial W^D(t;h)}{\partial t}$$

which takes its lowest value when $\gamma \to 0.5$, thus being greater than

$$\begin{array}{l} -\frac{\partial W^D(t;l)}{\partial t}\left[W^D\left(t_l^{\max};l\right) - W^D_A(l)\right]\left[W^D(t;h) - W^D_A(h)\right] \quad + \\ \left[W^D\left(t_l^{\max};l\right) - W^D_A(l)\right]\left[W^D(t;l) - W^D_A(l)\right]\frac{\partial W^D(t;h)}{\partial t}, \end{array}$$

which is strictly greater than 0 for $\tau \in \left[\tau_h^{min}, \tau_h^{max}\right]$ if the inequality condition in (6) holds. These inequalities imply that the first order condition associated with (4) hold with a strict inequality for $\tau \in \left[\tau_h^{min}, \tau_h^{max}\right]$, which in turn implies that $\tau_b = \tau_h^{max}$. C4 in the proof for Lemma 4 completes the proof with $\tau_b = \tau_h^{max}$.

A.4 Proof of Proposition 2

(i) The probability of litigation against an imposter is equal to $1 - \gamma\beta_l(\tau_h^{max}) - (1-\gamma)\beta_h(\tau_h^{max})$, which strictly decreases in γ with $\beta_l(\tau_h^{max}) > \beta_h(\tau_h^{max})$.

(ii) $\partial\beta_h(\tau_h^{max})/\partial\gamma < 0$ and $\partial\beta_l(\tau_h^{max})/\partial\gamma < 0$.

(iii) The probability of litigation against D_h's proposing τ_h^{max}, $\gamma\beta_h(\tau_h^{max}) + (1-\gamma)\beta_l(\tau_h^{max})$, strictly decreases in in γ with $\beta_l(\tau_h^{max}) > \beta_h(\tau_h^{max})$.

(iv) The expected payoff of D_l is equal to $W^D\left(t_l^{\max};l\right)$ to have (3) be satisfied with an equality. The expected payoff of C is equal to $\rho W^C_A(h) + (1-\rho)W^C_A(l)$ given the equilibrium strategy.

A.5 Proof of Lemma 7

Consider a pooling PBE under which both types of D propose τ_{pool} and C accepts it. Now consider an alternative tariff, τ', such that $\tau_{pool} < \tau' < \tau_h^{\max}$. C's strategies in response to the deviation is a probability of litigation. The range of litigation probabilities that makes deviation to τ' suboptimal for D_h is a subset of the corresponding range for D_l. Therefore, C will infer that the off-equilibrium action is likely taken by a high type, in which case C will not litigate (since $\tau' < \tau_h^{\max}$). Therefore, the pooling equilibrium is not Universally Divine.

We now show that partially-separating PBEs are not Divine either. Consider a partially-separating PBE in which D_l randomizes between $\tau_l^{\max}$ and $\tau_1 > \tau_h^{\min}$, and D_h chooses τ_1 with certainty. Now suppose that D_h defects to $\tau' \in (\tau, \tau_h^{\max}]$ and let β' denote the probability of litigation that makes D_l indifferent between $\tau_l^{\max}$ and τ'. This defection will strictly reduce D_l's expected welfare if C responds with a litigation probability $\beta > \beta'$. However, there is clearly $\beta'' > \beta'$ such that this defection will strictly increases D_h's expected welfare if C responds with any $\beta \in (\beta', \beta'')$. Therefore, if a defection to τ' is observed (instead of observing one of possible equilibrium outcomes, τ or $\tau_l^{\max}$) , C will believe it is committed by D_h. The standard argument, then, implies that the partially-separating equilibrium is not Divine. QED

A.6 Proof of Proposition 3

According to Lemma 7, we focus on a separating PBE that maximizes the expected payoff of D_h. In proving (i) D_l always proposes τ_l^{max} for settlement, which will be accepted by the complainant, one can apply the same logic as the one used for proving the same result in C1 for Lemma 4. Then, it remains to prove that (ii) results from solving the following constrained maximization problem for D_h:

$$\max_{\tau^S} \left\{ \beta W_A^D(h) + (1-\beta) W^D\left(\tau^S; h\right) \right\} \tag{17}$$

subject to (7) that holds with equality and $\beta_l = \beta_h = \beta$.

Due to the same reason for C2 to hold in the proof for Lemma 4, we can focus on $\tau^S \in \left[\tau_h^{min}, \tau_h^{max}\right]$.

Then, the first order condition associated with (17) is

$$\begin{aligned} -\tfrac{\partial \beta(\tau)}{\partial \tau}\left[W^D(\tau;h) - W_A^D(h)\right] \quad & + \\ [1-\beta(\tau)]\tfrac{\partial W^D(\tau;h)}{\partial \tau} \qquad & \geq 0, \end{aligned}$$

where

$$\beta(\tau) = \frac{W^D(\tau;l) - W^D\left(\tau_l^{\max};l\right)}{\left[W^D(\tau;l) - W_A^D(l)\right]},$$
$$\frac{\partial\beta(\tau)}{\partial\tau} = \frac{\frac{\partial W^D(\tau;l)}{\partial\tau}\left[W^D\left(\tau_l^{\max};l\right) - W_A^D(l)\right]}{\left[W^D(\tau;l) - W_A^D(l)\right]^2}.$$

This first order condition can be rewritten into:

$$\begin{array}{cc} -\frac{\partial W^D(\tau;l)}{\partial\tau}\left[W^D(\tau;h) - W_A^D(h)\right] & + \\ \left[W^D(\tau;l) - W_A^D(l)\right]\frac{\partial W^D(\tau;h)}{\partial\tau} & \geq 0 \end{array},$$

which holds with a strict inequality for $\tau \in \left[\tau_h^{min}, \tau_h^{max}\right]$ given the inequality condition in (6). As a solution to (17), thus we have $\tau^S = \tau_h^{max}$ and $\beta = \beta\left(\tau_h^{max}\right)$.

A.7 Proof of Proposition 4

First note that regardless of his information, C's settlement proposal, τ^S, must be either $\tau_l^{\min}$ or $\tau_h^{\min}$ because the response of either type of D will be identical for any offer between $\left(\tau_l^{\min}, \tau_h^{min}\right)$, and C will always prefer these extreme offers to any point in the middle.

If $\tau^S = \tau_l^{\min}$, then D_l will settle and D_h will litigate. Therefore, given θ^C, C's expected payoff from proposing $\tau^S = \tau_l^{\min}$ is

$$\Pr(\theta = h|\theta^C)W_A^C(h) + \left[1 - \Pr(\theta = h|\theta^C)\right]W^C\left(\tau_l^{\min}\right).$$

If $\tau^S = \tau_h^{\min}$, then both types of D will accept the proposal in which case the payoff of C is given by $W^C\left(\tau_h^{\min}\right)$. Therefore, C will propose $\tau^S = \tau_h^{\min}$ if and only if

$$W^C\left(\tau_h^{\min}\right) \geq \Pr(\theta = h|\theta^C)W_A^C(h) + \left[1 - \Pr(\theta = h|\theta^C)\right]W^C\left(\tau_l^{\min}\right),$$

or, equivalently, iff

$$\Pr(\theta = h|\theta^C) \geq \frac{W^C\left(\tau_l^{\min}\right) - W^C\left(\tau_h^{\min}\right)}{W^C\left(\tau_l^{\min}\right) - W_A^C(h)}.$$

A.8 Proof of Proposition 7

We define γ^I, γ^{II}, and γ^{III} as follows:

$\gamma^I \equiv [W^D(t_h^b;l) - W^D(t_l^C;l)]/[W^D(t_b^C;l) - W_A^D(l)]$, with t_h^b being implicitly defined by t solving

$$\frac{W^D(t;l) - W_A^D(l)}{W^D(t;h) - W_A^D(h)} = \frac{\frac{\partial W^D(t;l)}{\partial t}}{\frac{\partial W^D(t;h)}{\partial t}},$$

γ^{II} is the minimum value of γ that satisfies the following equality;

$$\frac{(1-\gamma)}{\gamma^2} = \frac{\frac{\partial W^D\left(\hat{t}_h^S(\gamma);h\right)}{\partial t}}{\frac{\partial W^D\left(\hat{t}_h^S(\gamma);l\right)}{\partial t}} \frac{\left[W^D\left(\hat{t}_h^S(\gamma);l\right) - W_A^D(l)\right]^2}{[W^D(t_l^C;l) - W_A^D(l)]\left[W^D\left(\hat{t}_h^S(\gamma);h\right) - W_A^D(h)\right]},$$

with $\hat{t}_h^S(\gamma)$ being defined by t that induces $\beta_h(t) = 0$ and $\beta_l(t) = 1$ using $D_l's$ incentive constraint in (10) that holds with equality;

$\gamma^{III} \equiv [W^D(t_h^C;l) - W^D(t_l^C;l)]/[W^D(t_h^C;l) - W_A^D(l)] < 1$.

With regard to relative size of γ^I, γ^{II}, and γ^{III}, $t_h^b > t_h^C$ implies that $\gamma^{III} > \gamma^I$, and the following proof for (ii) also shows that $\gamma^{II} \in (\gamma^I, \gamma^{III})$.

(i) We show this result in the following two steps: first, assume that the equilibrium arbitration strategy profile embodies $\beta_l\left(t_h^S\right) = 1$ and $\beta_h\left(t_h^S\right) > 0$, establishing the associated comparative static results; second, the equilibrium arbitration strategy profile embodies $\beta_l\left(t_h^S\right) = 1$ and $\beta_h\left(t_h^S\right) > 0$, if $\gamma < \gamma^I$. If the Divine PBE embodies $\beta_l\left(t_h^S\right) = 1$ and $\beta_h\left(t_h^S\right) > 0$, it implies that the first order derivative associated with the maximization problem in (10) takes the following form and equals to zero:

$$-\gamma\frac{\partial \beta_h(t)}{\partial t}\left[W^D(t;h) - W_A^D(h)\right] + \gamma[1-\beta_h(t)]\frac{\partial W^D(t;h)}{\partial t} = 0. \tag{18}$$

with

$$\begin{aligned}
\frac{\partial \beta_h(t)}{\partial t} &= \frac{1}{(1-\gamma)} \frac{\partial W^D(t;l)}{\partial t} \frac{\left[W^D\left(t_l^C;l\right) - W_A^D(l)\right]}{\left[W^D(t;l) - W_A^D(l)\right]^2}, \\
1-\beta_h(t) &= \frac{1}{(1-\gamma)} \frac{\left[W^D\left(t_l^C;l\right) - W_A^D(l)\right]}{\left[W^D(t;l) - W_A^D(l)\right]}, \\
\frac{\partial \beta_h(t)}{\partial \gamma} &= -\frac{\left[W^D\left(t_l^C;l\right) - W_A^D(l)\right]}{(1-\gamma)^2\left[W^D(t;l) - W_A^D(l)\right]} < 0.
\end{aligned} \tag{19}$$

The first two equalities in (19) implies that (18) is not affected by γ, which in turn implies that $\frac{\partial t_h^S}{\partial \gamma} = 0$. In fact, (18) implicitly define t_h^b as defined above, having $t_h^S = t_h^b$. Given this result, the last inequality in (19) implies $\frac{\partial \beta_h(t_h^S)}{\partial \gamma} < 0$. It remains to show that the Divine PBE embodies $\beta_l\left(t_h^S\right) = 1$ and $\beta_h\left(t_h^S\right) > 0$ if $\gamma < \gamma^I$. If $\gamma = \gamma^I$, then $\beta_l(t^S) = 1$ and $\beta_h(t^S) = 0$ to make D_l be indifferent between proposing t_l^{max} and proposing $t_h^S = t_h^b$, which in turn implies that $\beta_l(t_h^S) = 1$ and $\beta_h(t_h^S) > 0$ are required to make D_l be indifferent between proposing t_l^C and $t_h^S = t_h^b$ if $\gamma < \gamma^I$.

(ii) Given $\gamma = \gamma^I$, the first order derivative associated with the maximization problem in (10) takes the following form, being strictly smaller than zero for $t \geq t_h^b$ in the neighborhood of t_h^b if $\gamma^I > 0.5$:

$$-(1-\gamma)\frac{\partial \beta_l(t)}{\partial t}\left[W^D(t;h) - W_A^D(h)\right] + \left[1-(1-\gamma)\beta_l(t)\right]\frac{\partial W^D(t;h)}{\partial t} < 0. \tag{20}$$

with

$$\begin{aligned}
\frac{\partial \beta_l(t)}{\partial t} &= \frac{1}{\gamma} \frac{\partial W^D(t;l)}{\partial t} \frac{\left[W^D\left(t_l^C;l\right) - W_A^D(l)\right]}{\left[W^D(t;l) - W_A^D(l)\right]^2} < 0, \\
1-(1-\gamma)\beta_l(t) &= 1 - \frac{(1-\gamma)}{\gamma} \frac{\left[W^D(t;l) - W^D\left(t_l^C;l\right)\right]}{\left[W^D(t;l) - W_A^D(l)\right]}, \\
\frac{\partial \beta_l(t)}{\partial \gamma} &= -\frac{\left[W^D(t;l) - W^D\left(t_l^C;l\right)\right]}{\gamma^2\left[W^D(t;l) - W_A^D(l)\right]} < 0.
\end{aligned} \tag{21}$$

First, note that D_l's incentive constraint in (10) requires $\beta_h(t_h^b) = 0$ and $\beta_l(t_h^b) = 1$ by definition of γ^I, and $\beta_h(t) = 0$ and $\beta_l(t) < 1$ for $t > t_h^b$. The strict inequality in (20) for $t > t_h^b$ comes from the definition of t_h^b (being the unique value of t that maximizes the expected payoff of D_h with $\gamma = \gamma^I$). The strict inequality in (20) holds even at $t = t_h^b$ because the first term in (20) is strictly smaller than the first term in (18) at $t = t_h^b$ with $\gamma = \gamma^I > 0.5$.

This strict inequality in (20) at $t = t_h^b$ implies $\beta_l(t_h^S) = 1$ and $\beta_h(t_h^S) = 0$ with $\frac{\partial \beta_l(t_h^S)}{\partial \gamma} = 0$ for $\gamma > \gamma^I$ in the neighborhood of γ^I. On the one hand, an increase in γ raises the absolute value of the second term both in (18) and (20), thus raising the negative effect from raising t in maximizing the expected payoff of D_h. An increase in γ also decreases $\hat{t}_h^S(\gamma)$ with $\partial \hat{t}_h^S / \partial \gamma < 0$. On the other hand, an increase in γ decreases the first term in (20) but increases the first term in (18). These comparative static results implies that a small increase in γ at $\gamma = \gamma^I$ will lead to a decrease in t_h^S (thus, $\partial t_h^S / \partial \gamma < 0$) so that $t_h^S = \hat{t}_h^S(\gamma)$. This is because the first order derivative associated with the maximization problem in (10) is positive for $t_h^S < \hat{t}_h^S(\gamma)$ and it is negative for $t_h^S > \hat{t}_h^S(\gamma)$. For $\gamma \geq \gamma^I$, therefore, $\beta_l\left(t_h^S\right) = 1$ and $\beta_h\left(t_h^S\right) = 0$ with $\partial t_h^S / \partial \gamma < 0$ and $\frac{\partial \beta_l(t_h^S)}{\partial \gamma} = 0$ in the neighborhood of γ^I.

Recall that we define γ^{II} as the minimum value of γ such that the left hand side of the inequality in (20) becomes zero with $t_h^S = \hat{t}_h^S(\gamma)$, thus $\gamma^{II} > \gamma^I$. Note that such a value of γ (denoted by γ^{II}) exists and it is strictly smaller than γ^{III}. Because $\hat{t}_h^S(\gamma) = t_h^C$ if $\gamma = \gamma^{III}$, having the left hand side of the inequality in (20) is greater than zero at $t = \hat{t}_h^S(\gamma) = t_h^C$, γ^{II} does exist and it is smaller than γ^{III}.

(iii) If $\gamma \geq \gamma^{II}$, then $\beta_h\left(t_h^S\right) = 0$ because the first order derivative associated with the maximization problem in (10) is strictly greater than zero for any $t < \hat{t}_h^S(\gamma)$. This implies that $\beta_l\left(t_h^S\right) = 1$ with $t_h^S = \hat{t}_h^S(\gamma)$ or $\beta_l\left(t_h^S\right) < 1$ with $t_h^S > \hat{t}_h^S(\gamma)$ if $\gamma \geq \gamma^{II}$, which in turn implies that $\partial t_h^S / \partial \gamma < 0$: $\partial t_h^S / \partial \gamma < 0$ for the case in which $\beta_l\left(t_h^S\right) < 1$ is shown in the following proof for (iv).

(iv) If $\gamma \geq \gamma^{III}$, first note that $\beta_h(t) = 0$ and $\beta_l(t) < 1$ under the requirement of D_l's incentive constraint in (10) for all $t \in T^C\left(t_h^C\right)$ with $t_h^S > t_h^C$ by the definition of γ^{III}. This implies that the first derivative associated with the maximization problem in (10) takes the same form as the one in (20), and it equals to zero at t_h^S. When γ increases, the first term in (20) decreases, reflecting a reduced benefit from raising t, and the (negative) second term in (20) decreases, reflecting an increased cost from raising t. These changes in (20) in response to an increase in γ implies $\partial t_h^S / \partial \gamma < 0$. To show that $\partial t_h^S / \partial \gamma < 0$ for any Divine PBE with $\beta_l\left(t_h^S\right) \in (0,1)$ and $\beta_h\left(t_b\right) = 0$, we can apply the same logic. This is because the first derivative associated with the maximization problem in (10) takes the same form as the one in (20) in the neighborhood of $t = t_h^S$ if $\beta_l\left(t_h^S\right) \in (0,1)$ and $\beta_h\left(t_h^S\right) = 0$.

$\lim_{\gamma \to 1} (t_h^S) = t_h^C$ because the limit of the firm term in (20) approaches to zero with $\gamma \to 1$ while the second term in (20) remains negative, except at $t = t_h^C$. Finally, D_l's incentive constraint in (10) implies $\beta_l(t) = \left[W^D\left(t_h^S; l\right) - W^D\left(t_l^C; l\right)\right] / \gamma \left[W^D\left(t_h^S; l\right) - W_A^D(l)\right]$ for $\beta_l(t) < 1$, thus

$$\lim_{\gamma \to 1} \beta_l\left(t_h^S\right) = \frac{W^D\left(t_h^C; l\right) - W^D\left(t_l^C; l\right)}{W^D\left(t_h^C; l\right) - W_A^D(l)} > 0.$$

References

Ahn, D., J. Lee, and J.-H. Park (2014). Understanding non-litigated disputes in the wto dispute settlement system. *Journal of World Trade 47*(5), 985–1012.

Amador, M. and K. Bagwell (2013). The Theory of Optimal Delegation With an Application to Tariff Caps. *Econometrica 81*(4).

Ayres, I. and B. J. Nalebuff (1996). Common knowledge as a barrier to negotiation. *UCLA L. Rev. 44*, 1631.

Bagwell, K. (1995). Commitment and observability in games. *Games and Economic Behavior 8*(2), 271–280.

Bagwell, K. (2009). Self-enforcing Trade Agreements and Private Information. *NBER Working Paper*.

Bagwell, K. and R. Staiger (2005). Enforcement, Private Political Pressure, and the General Agreement on Tariffs and Trade/World Trade Organization Escape Clause. *The Journal of Legal Studies 34*(2), 471–513.

Banks, J. S. and J. Sobel (1987). Equilibrium selection in signaling games. *Econometrica 55*(3), 647–661.

Bergemann, D. and S. Morris (2009). Robust implementation in direct mechanisms. *The Review of Economic Studies 76*(4), 1175–1204.

Beshkar, M. (2010a). Optimal Remedies in International Trade Agreements. *European Economic Review 54*(3), 455 – 466.

Beshkar, M. (2010b). Trade Skirmishes and Safeguards: A Theory of the WTO Dispute Settlement Process. *Journal of International Economics 82*(1), 35 – 48.

Beshkar, M. (2016). Arbitration and renegotiation in trade agreements. *Journal of Law, Economics, and Organization 32*(3), 586–619.

Beshkar, M. and E. Bond (2017). Cap and Escape in Trade Agreements. *American Economic Journal-Microeconomics 9*(4), 171–202.

Beshkar, M., E. Bond, and Y. Rho (2015). Tariff Binding and Overhang: Theory and Evidence. *Journal of International Economics 97*(1), 1–13.

Beshkar, M. and E. W. Bond (2016). The escape clause in trade agreements. In K. Bagwell and R. W. Staiger (Eds.), *Handbook of Commercial Policy*, Volume B1, pp. 69–106. Elsevier.

Beshkar, M. and M. Majbouri (2019). Estimating a Model of Settlement Bargaining in the World Trade Organization. In B. Zissimos (Ed.), *The WTO and Economic Development*. MIT Press.

Beshkar, M. and J. Park (2017). Dispute Settlement with Second-Order Uncertainty: The Case of International Trade Disputes. *Working Paper*.

Bown, C. P. (2004). On the economic success of GATT/WTO dispute settlement. *Review of Economics and Statistics 86*(3), 811–823.

Bown, C. P. and K. M. Reynolds (2015). Trade flows and trade disputes. *The Review of International Organizations 10*(2), 145–177.

Bown, C. P. and K. M. Reynolds (2017). Trade agreements and enforcement: evidence from wto dispute settlement. *American Economic Journal: Policy 9*(4), 64–100.

Chen, Y.-C., A. Di Tillio, E. Faingold, and S. Xiong (2017). Characterizing the strategic impact of misspecified beliefs. *The Review of Economic Studies 84*(4), 1424–1471.

Daughety, A. F. and J. F. Reinganum (1994). Settlement negotiations with two-sided asymmetric information: Model duality, information distribution, and efficiency. *International Review of Law and Economics 14*(3), 283–298.

Daughety, A. F. and J. F. Reinganum (2017). Settlement and trial. In F. Parisi (Ed.), *Handbook of Law and Economics*, Volume III, pp. 229–246. Oxford University Press.

Feinberg, Y. and A. Skrzypacz (2005). Uncertainty about uncertainty and delay in bargaining. *Econometrica 73*(1), 69–91.

Hörner, J., M. Morelli, and F. Squintani (2015). Mediation and peace. *The Review of Economic Studies 82*(4), 1483–1501.

Kuenzel, D. J. (2017). WTO dispute determinants. *European Economic Review 91*, 157–179.

Maggi, G. (1999). The value of commitment with imperfect observability and private information. *The RAND Journal of Economics*, 555–574.

Maggi, G. and R. Staiger (2011). The role of dispute settlement procedures in international trade agreements. *The Quarterly Journal of Economics 126*(1), 475–515.

Maggi, G. and R. W. Staiger (2017). Learning by ruling and trade disputes. Working paper, NBER.

Maggi, G. and R. W. Staiger (2018). Trade disputes and settlement. *International Economic Review 59*(1), 19–50.

Morris, S. and H. S. Shin (2002). Social value of public information. *The American Economic Review 92*(5), 1521–1534.

Morris, S., H. S. Shin, and M. Yildiz (2016). Common belief foundations of global games. *Journal of Economic Theory 163*, 826–848.

Park, J.-H. (2011). Enforcing international trade agreements with imperfect private monitoring. *The Review of Economic Studies 78*(3), 1102–1134.

Park, J.-H. (2016). Enforcement and dispute settlement. In K. Bagwell and R. W. Staiger (Eds.), *Handbook of Commercial Policy*, Volume B1, pp. 3–67. Elsevier.

Schweizer, U. (1989). Litigation and settlement under two-sided incomplete information. *The Review of Economic Studies 56*(2), 163–177.

Spier, K. E. (2007). Litigation. In A. M. Polinsky and S. Steven (Eds.), *Handbook of Law and Economics*, Volume 1, pp. 259–342. North Holland.

Sykes, A. (2003). The Safeguards Mess: A Critique of WTO Jurisprudence. *World Trade Review 2*(03), 261–295.

Sykes, A. O. (2016). Legal aspects of commercial policy rules. In K. Bagwell and R. W. Staiger (Eds.), *Handbook of Commercial Policy*, Volume A1, pp. 263–332. Elsevier.

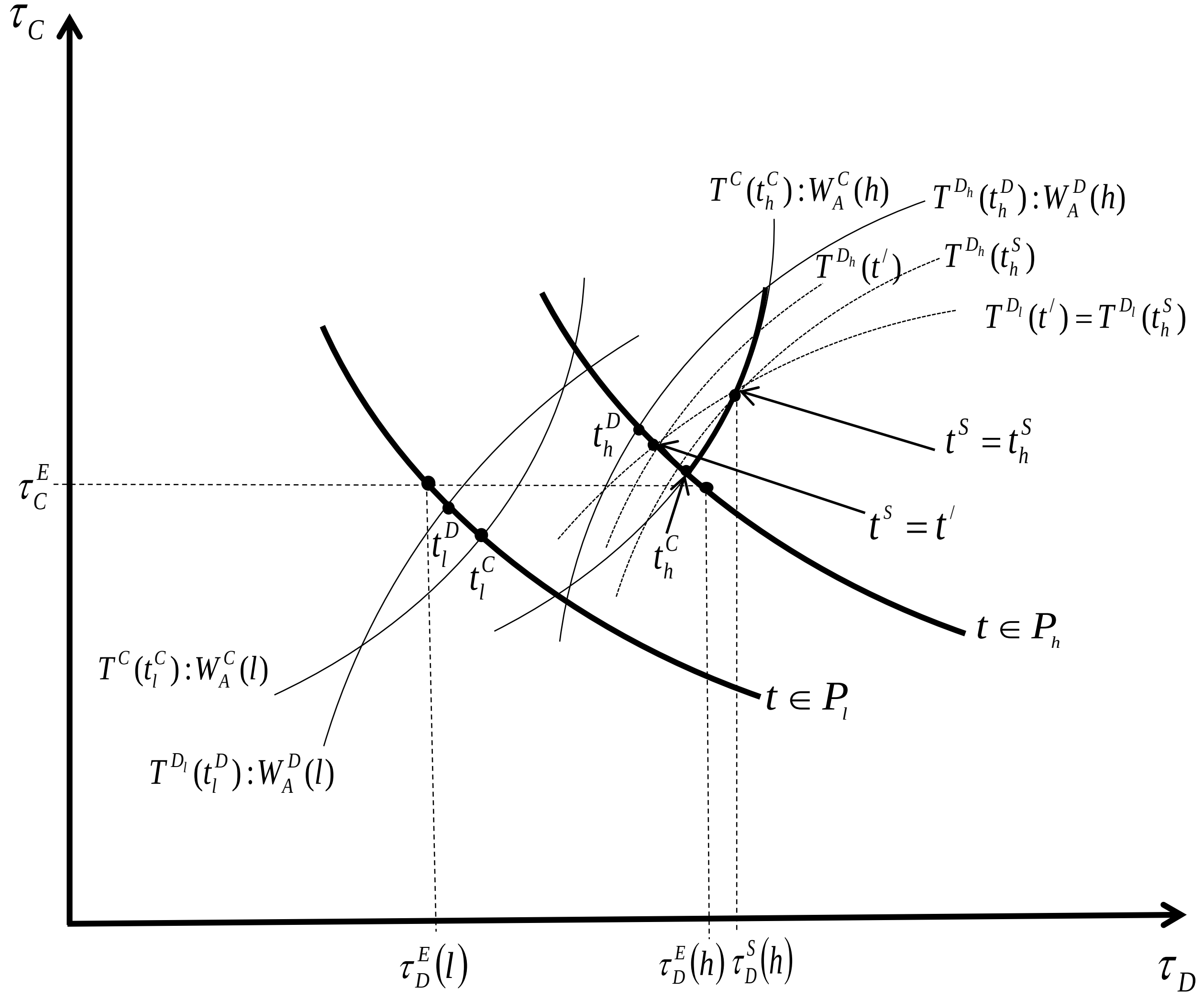

Figure 1. Potential Settlement Gains and Equilibrium Settlement Offer

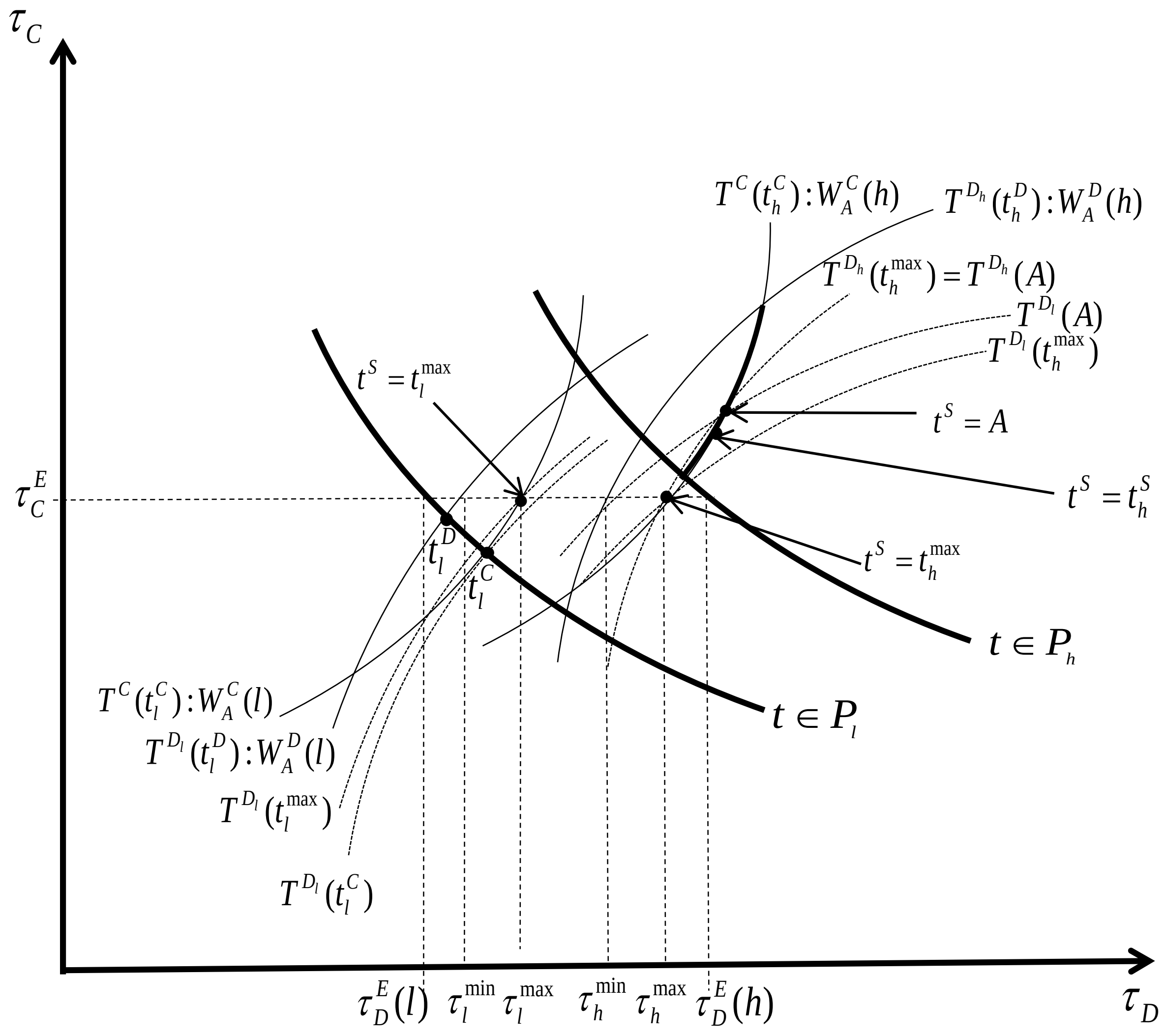

Figure 2. Comparison of One Variable Case with Two Variable One

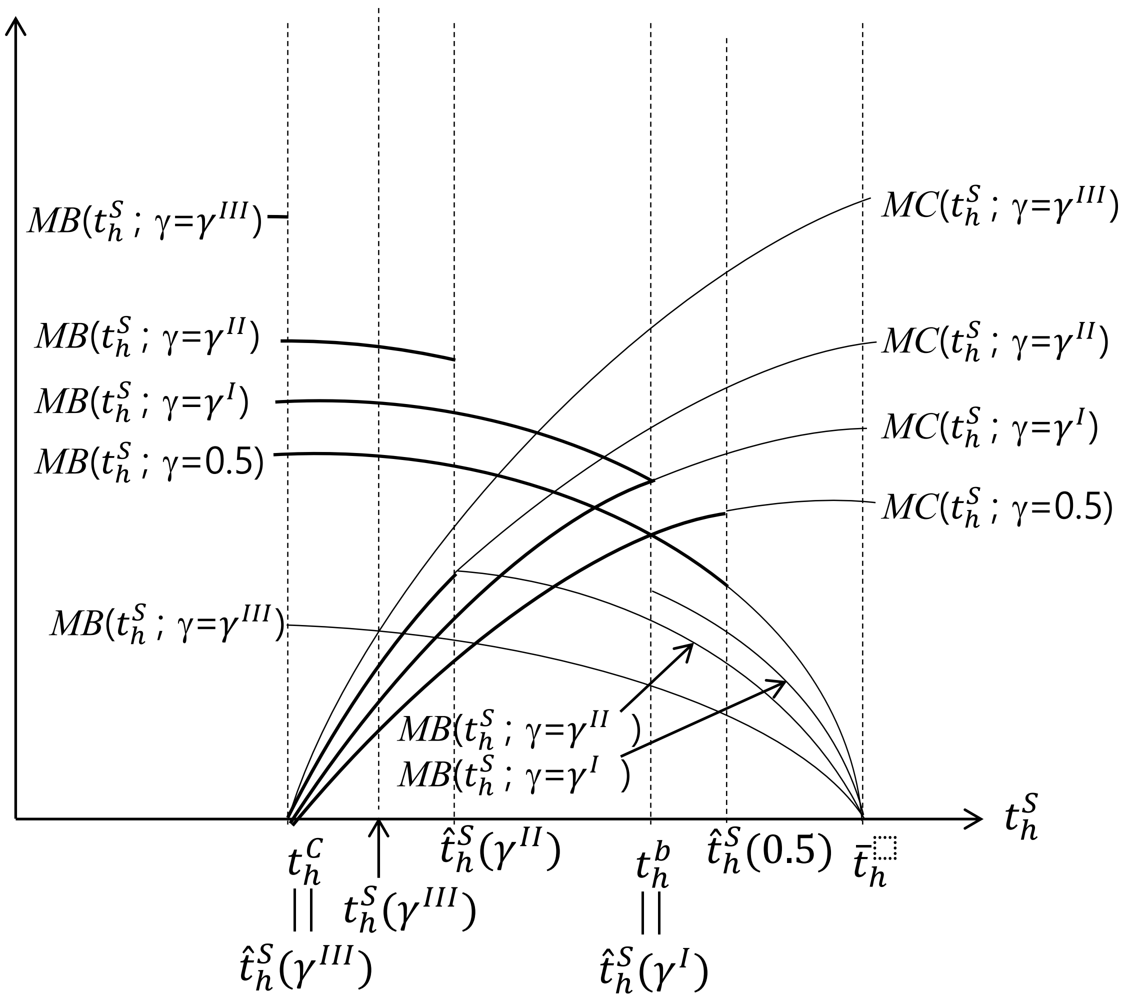

Figure 3. Effect of γ on the Equilibrium Settlement Proposal

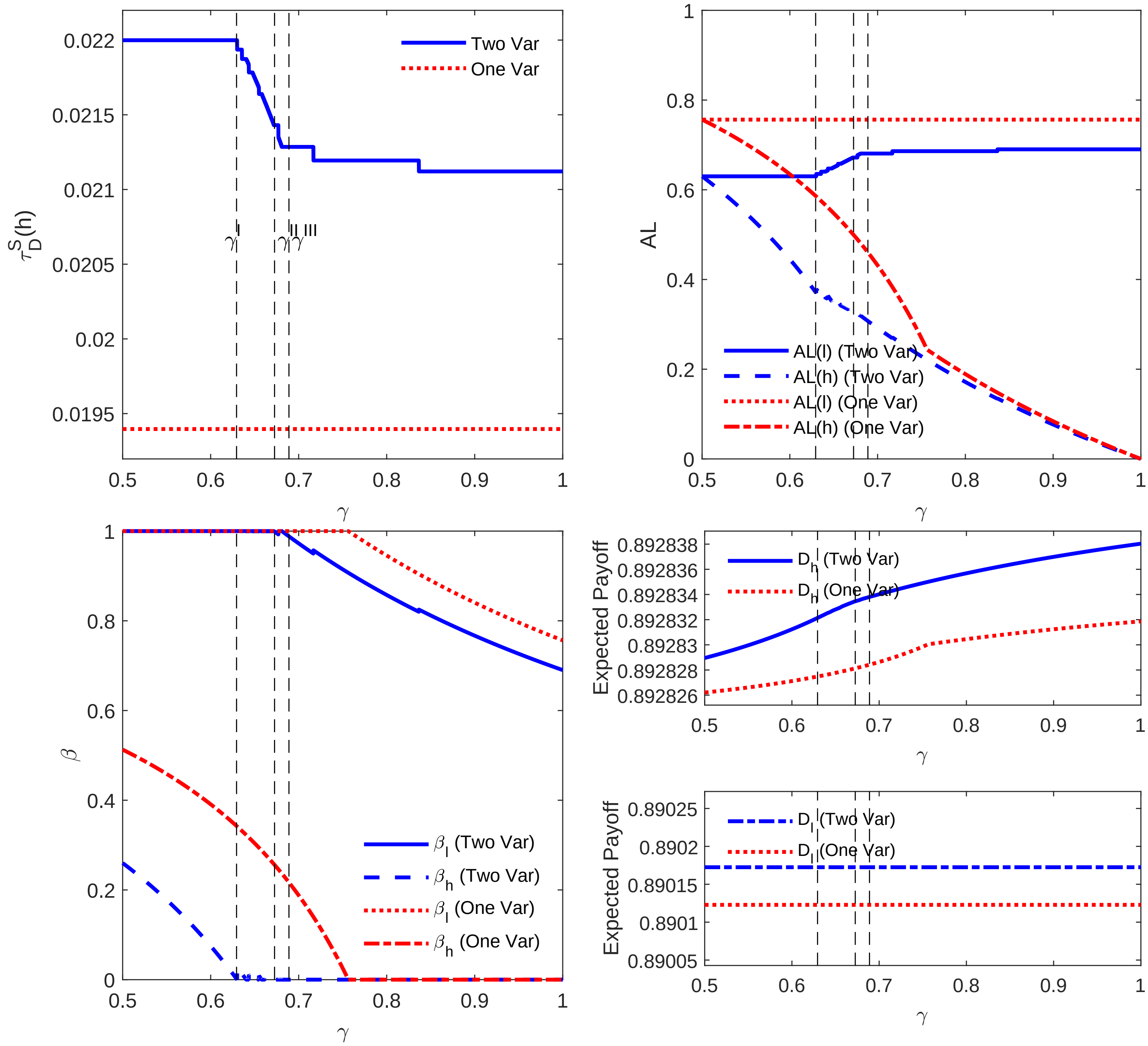

Figure 4. Numerical Analysis Based on an International Trade Model