Dynamics of Risky Agreements

Renee Bowen* Malte Lammert† Aleksandr Levkun‡

November 24, 2025

Abstract

We investigate the efficiency of agreements with the following features: (i) *self-enforcing*—any agent can walk away from the agreement at any moment; (ii) *dynamic*—payouts occur stochastically while the agreement is in force; (iii) *risky*—one agent is more favored by the agreement but the favored agent is unknown ex-ante. These features appear in international economic agreements, such as the WTO. Such arrangements have formal or informal mechanisms to resolve disputes that may be more favorable to one agent, but who is favored is learned only as disputes arise. Welfare combines the agents' joint surplus with a positive societal externality generated while the agreement is in force. Some efficient agreements never form, and risky agreements end with certainty as disputes provide information about favor status. We find that when the agreement stakes are high, welfare is maximal at intermediate beliefs that balance optimism sufficient for agents' participation with the sustained uncertainty that makes the agreement socially valuable. Slow judgments are optimal when the externality is sufficiently strong. While global welfare is maximized by symmetric judgment frequencies, mild asymmetry can locally increase welfare by promoting participation and improving the informational value of the agreement.

*Renee Bowen, Georgetown University and NBER. † Malte Lammert, PwC. ‡ Aleksandr Levkun, WU (Vienna University of Economics and Business).

We thank S. Nageeb Ali, Emin Karagözoğlu, Joel Sobel, Joel Watson, and seminar participants at Bilkent University, Columbia University, Georgetown University, MIT, Michigan State University, Penn State University, University of Houston, University of Pittsburgh, Vanderbilt University, and Yale University for useful comments and suggestions.

The Devil is in the details.

1 Introduction

While features of contracts have been studied at length, the extent to which agreements interact with the system of handling disputes is less well understood. Agreements—such as constitutions, international agreements, insurance contracts, or entrepreneurial ventures—rely on mechanisms to resolve disputes that periodically arise between parties. Domestic courts often play this role, however, international agreements typically include their own dispute resolution process codified in the agreement. A feature of such dispute systems is that they may favor one party or another. For example, a plaintiff or defendant with more financial resources may be expected to navigate the courts more successfully with better lawyers. Or the insured, as a consumer, may be presumed at an advantage in an insurance dispute. Larger countries with more economic or military might may anticipate favorable treatment in international disputes. Potential bias in the dispute system can surface only when the agreement provides no clear guidance, i.e., when there is ambiguity. With ambiguity, the dispute system can only rely on its judgment. For any particular agreement and dispute system, if bias does exist, it is revealed over time as disputes arise and are resolved more or less in one party's favor. As a party learns who is favored, they may be prompted to exit the agreement because their expected value is diminished. In the case of a constitution, elites or citizens may choose to alter the constitution or abandon it in favor of a new one. In the case of an insurance contract, the insured may switch carriers. International agreements often end with parties walking away from their obligations. The possibility that a party can be less favored by an agreement makes the agreement *risky*. In this paper we take a first step to study such risky agreements, their welfare properties, and optimal design.

A prominent example of the formation and dissolution of a risky agreement is the World Trade Organization's Appellate Body (AB). Established in 1995 as part of the Marakesh Agreement, the AB was meant to act as a "final court" to arbitrate disputes between WTO member countries. The challenge was that the agreement gave little specificity on how the AB should rule on some highly sensitive cases. This ambiguity became a flash point related to the US frequent use of trade remedies—tariff protection allowed by the WTO in special cases. While the trade remedies were allowed by the WTO, what circumstances qualified for trade remedies was left ambiguous. As the WTO informally adopted an initial judgment against the US as precedent, they established a bias against the US for future cases. The US learned over time that the AB's judgments in practice were unfavorable to its use of trade remedies and disabled it in 2019 citing consistent losses due to "judicial overreach".

Another example of a risky agreement is the US-Mexico-Canada Agreement (USMCA). Signed in July 2020 to replace the North-American-Free-Trade-Agreement (NAFTA), it was hailed as a win by the United States.[1] In January 2023, however, the United States lost its biggest case under USMCA related to a disagreement about the calculation of the share of manufactured autos originating in a member country. A Canadian newspaper, CBC, reported *"Canada, Mexico and auto companies have been declared the winners in arguably the most important trade dispute under the new NAFTA, landing the U.S. on the losing side"*(CBC, 2023). The larger disagreement between the United States, Mexico and Canada was about the interpretation of the rules regarding calculation of the shares. CBC News reported that *"Canada submitted an email sent by a U.S. official that supported the complainants' claim that all three countries originally understood they were agreeing to the simpler formula"*.

Risky agreements depend critically on the agents' beliefs about the frequency of rewards generated by the agreement. A reward can be thought of as a positive judgment in case of a dispute when the agreement language was unclear. The formation of an agreement requires sufficient optimism from both agents about how such judgments are resolved, as any agent will reject the agreement if they are certain that judgments will never be resolved in their favor. Once in force, the agreement will end if at least one agent becomes sufficiently pessimistic about how these judgments will be handled. In this paper we ask: 1) Under what conditions are risky agreements established? 2) Under what conditions do they break down? 3) How do agreement features, such as the frequency of judgments, impact the welfare consequences of risky agreements?

We present a model of an agreement between two agents with the following features. The agreement is: (i) *risky*, i.e. the distribution of rewards is uncertain; (ii) *self-enforcing*, i.e. either agent can quit at any time without penalty; and (iii) *dynamic*, i.e. returns arrive stochastically while the agreement is in force. In addition to international agreements, constitutions, entrepreneurial ventures, co-authoring research, options agreements, insurance agreements (disputes over what is covered or not) fall into this category. We assume that only one agent is favored by the agreement and that the favored agent is ex-ante unknown, but can be learned over time.

Formally, we consider that each agent has a risky arm of a Poisson bandit that delivers lump sum rewards with some frequency, but also has access to a safe arm that delivers a steady flow of lower payoffs, as in Keller and Rady (2010) and Keller and Rady (2015). The risky arms are either good—delivering rewards with a high frequency—or bad—delivery rewards with a low frequency. The frequency of rewards on the risky arm can be understood as the frequency of judgments in an agent's favor. Agents may have different frequencies of judgment and we consider agent 1 to be the agent with the higher frequency of judgments conditional on having a good arm. Exactly one agent has the good arm and one agent has the bad arm, while the agent with the good arm is ex-ante unknown. The bandit arms are thus negatively correlated as in

[1] The office of the US Trade Representative highlighted it as *"creating a more level playing field for American workers, including improved rules of origin for automobiles, trucks, other products, and disciplines on currency manipulation"* (USTR, 2020).

Klein and Rady (2011), so that a good risky arm for one agent implies a bad risky arm for the other. We study the evolution of the belief that agent 1 has the good risky arm.[2] A high belief is good for agent 1 and a low belief is good for agent 2. We assume, for simplicity, that a bad risky arm never yields a reward, so a return for any agent resolves the uncertainty fully. In other words, we assume that the news is conclusive.

The agreement is in force if both agents pull their risky arm, and ends if at least one agent quits their arm reflecting the fact that any agent can terminate the agreement unilaterally. The novelty of our bandit structure is that experimentation requires both agents to simultaneously use their risky arms. If one agent switches to their safe arm, experimentation and learning halt for all agents. We study equilibria in cut-off strategies, in which agreements are in force for an interval of beliefs. That is, each agent must be sufficiently optimistic about their own arm to keep experimenting in equilibrium. This implies that some agreements are never formed, and all risky agreements end in the long-run. We study the cut-off beliefs to answer our research questions.

To build intuition we begin with the case of symmetric agents who have the same arrival rate of judgments, conditional on having a good arm. Counter to standard intuition in the bandit literature, symmetric agents behave myopically, and simply weigh the short term benefit of the safe arm against the instantaneous expected benefit of pulling the risky arm. The reason is that with symmetric agents, no learning occurs and hence there is no option value to remaining in the agreement. In other words, there is no value of experimenting. The myopic cutoff beliefs for symmetric agents are straightforward to calculate and facilitate studying the welfare implications in detail.

We consider welfare as a convex combination between the *agreement value* to the agents themselves, and the *societal value*, which can be thought of as a positive externality. In each of the trade agreement examples presented, the negative judgments affected a narrow constituency. We can think of these constituencies as the parties to the agreement, because they lobby or influence the government to remain in or leave the agreement. However, the agreements as a whole sustain lower tariffs, generating wider societal benefits from reduced prices and increased productivity. We do not take a stance on which value is more salient, and consider arbitrary convex combinations. In the symmetric case, the agreement value increases with the frequency of judgments, as agents prefer to receive the uncertain reward sooner. In contrast, the societal value decreases with the frequency of judgments, as the externality generated from the agreement continues as long as the agreement is in force. If sufficient weight is placed on societal value the optimal frequency of judgment is as slow as possible, without eliminating the agents' incentives to maintain the agreement. A longer agreement implies more societal benefit, and delayed judgments allow the uncertainty and the agreement to persist. These forces are also present in the asymmetric case, but another force arises due to the learning that occurs.

[2] The belief that agent 2 has the good risky arm is simply 1 minus agent 1's belief because of the perfect negative correlation.

If agents are asymmetric, then in the absence of news learning occurs and beliefs drift down. We find that the agent with the higher frequency of judgment, agent 1, still behaves myopically, but the agent with a lower frequency of judgments, agent 2, requires less optimism than a myopic agent to form the agreement. This is because the absence of the arrival of a reward is evidence in favor of agent 2. Agent 2's belief threshold increases monotonically with both agents' frequency of judgments, whereas agent 1's cutoff is only influenced by agent 1's frequency of judgments. We show that depending on the stakes of the agreement—the relative advantage of the good risky arm over the safe payoff—welfare can be either monotonically increasing or hump-shaped in the experimentation region of beliefs. When the agreement stakes are high and societal value is important for welfare, welfare is maximized at intermediate beliefs that balance the agents' optimism required for participation with the uncertainty that sustains societal value.

We demonstrate that globally, asymmetry in judgment frequencies reduces welfare because belief drift shortens agreements on average and introduces dynamic inefficiencies. However, locally, we show that a mild asymmetry between agents can improve welfare through two distinct channels: by expanding the range of starting beliefs that support participation, or by enhancing the informational value of the agreement. These findings highlight that while perfect symmetry maximizes welfare overall, modest asymmetry can be welfare-improving.

The remainder of this paper is structured as follows. Section 2 reviews related literature. Section 3 introduces the formal model, the equilibrium in cut-off beliefs, and welfare metric. Section 4 lays out the analysis for symmetric agents. Section 5 analyzes the case of asymmetric agents pointing out the role of asymmetry and belief drift. Section 6 concludes. All omitted proofs are in Appendix A.

2 Related Literature

Our analysis is centered around the role of beliefs in agreements. This has been studied by Eliaz and Spiegler (2006), Chiappori and Salanie (2013), Giat and Subramanian (2013), and others. We add to this literature by linking the formation, dissolution and duration of agreements to the evolution of beliefs using strategic experimentation with rational agents.

Our analysis is also closely related to the literature on strategic experimentation, as pioneered by Bolton and Harris (1999). A bandit is a probabilistic device that generates observable evidence in accordance to predetermined distributions, conditional on an unknown state. Through observations of the evidence (or news arriving), agents update their beliefs. Two major distinctions can be made: Firstly, when multiple agents operate multiple bandits, these bandits can be positively or negatively correlated. If bandits are positively correlated, typically free-riding problems occur. Keller et al. (2005) study a model of positive correlation and conclusive news, where evidence will only arrive in one state, so that the observation of such evidence reveals the state. Keller and Rady (2010) study positive correlation when news is inconclusive. If bandits are negatively correlated, the free-riding incentive is reversed and experimentation is encouraged. Klein and Rady (2011) study negatively correlated bandits in the environment where news is conclusive and experimentation is individual. They organize their findings within three categories, depending on the stakes, i.e., the size of potential gains from the risky arm relative to the safe payoff. They find that for low stakes, at most one player plays risky, so that in equilibrium players use their single-agent cutoff. For intermediate and high stakes, players may apply their myopic cutoffs. Our analysis also depends on stakes, however, takes a different from. We share the notion that with high or higher stakes, some agents act myopically. Relevant to our analysis, Keller and Rady (2015) study a model of positive correlation and inconclusive news where the news is more likely to occur in an undesired state, so that upon observation of news, beliefs jump downwards rather than upwards, leading to failure of smooth pasting condition also present in our framework. We study a model of negative correlation under asymmetric arrival rates and conclusive news with the novelty that if one agent stops experimenting, experimentation stops for all agents. This interdependence is crucial for modeling dynamic agreements: If one agent quits the agreement, the agreement ends and all learning about the distribution of returns stops immediately. The notion of collective experimentation arises in Strulovici (2010), where symmetric agents vote on whether to continue experimenting. Our setting can be viewed as collective experimentation with negatively correlated bandits under the unanimity rule, where the agreement generates social externalities. Our analysis highlights the distinct role that asymmetry plays for welfare.

Our analysis further relates to the literature on ambiguity in agreements. Closest here is Bernheim and Whinston (1998) who argue that agreements may be kept uncertain for strategic reasons. Specifically, the uncertainty about some elements of the agreement may encourage a desired behavior of the opposite party in a dimension that is easily observable but not verifiable. Our work further relates to Riedel and Sass (2013, 2014) who allow one player in a game to decide over the ambiguity perceived by other players (Ellsberg game). They find that if the other players are ambiguity averse, then the resulting Ellsberg equilibria may lie outside of the support of Nash equilibrium, so that ambiguity can be constructive in achieving outcomes that are not supported by Nash equilibrium. Our results are related as our analysis shows how sufficient uncertainty can be constructive in the formation of agreements that would not have formed without uncertainty. Grant et al. (2014) consider the optimality of liquidated damages agreements in a setting of agreement ambiguity and potential for disputes. They show that when parties are ambiguity averse enough, they will optimally choose liquidated damages agreements and sacrifice risk-sharing opportunities. Tillio et al. (2016) demonstrate that a seller can benefit from an ambiguous buying mechanism which hides certain features of the mechanism when the buyer is ambiguity averse. Grimmelmann (2019) shows that such ambiguity can be found in smart agreements written in programming languages. Dütting et al. (2024) show that an infusion of ambiguity into the design of contracts can benefit the principal, and this benefit can be arbitrarily high. While all of these results require ambiguity rather than risk and most of them some degree of ambiguity aversion, our results do not rely on ambiguity or risk aversion.

3 The Model

There are two agents, $i = 1, 2$. Time is continuous, $t \in [0, \infty)$. At each instant of time t agents choose to participate in an agreement or not, where $k_{i,t} = 1$ indicates that agent i participates at time t and $k_{i,t} = 0$ means agent i does not participate at time t. The agreement is in force at time t if and only if both agents participate, i.e., $k_{1,t} = k_{2,t} = 1$. The agreement is not in force at time t if at least one agent chooses not to participate, i.e, $k_{i,t} = 0$ for some i. There is no exogenous enforcement or punishment, so participation is entirely self-enforcing.

Payoffs We assume that when the agreement is not in force, each agent receives a certain flow payoff $s > 0$, but when the agreement is in force, payoffs are stochastic and depend on which agent is favored. Specifically, we assume that, while in force, the agreement is favorable to one agent and unfavorable to the other, but it is unknown who is favored by the agreement. If the agreement is favorable to agent i it yields a lump-sum reward $h > 0$ for agent i over time interval dt with probability $\lambda_i dt$, and yields no reward to agent $j \neq i$.[3] Both agents observe the arrival of a lump-sum reward for either agent. The arrival of a lump-sum reward for agent i is therefore conclusive news that agent i is favored by the agreement. We assume, without loss, that $\lambda_1 \geq \lambda_2$. That is, conditional on the agreement favoring agent 1, the rate of arrival of rewards is weakly higher than if the agreement favors agent 2. We interpret λ_i as a *judgment frequency* for agent i. We further assume that a utilitarian social planner prefers the agreement to be in force rather than no agreement, i.e., $\lambda_2 h > 2s$. This implies $\lambda_i h > s > 0$ for each agent i, so a favorable agreement is better than no agreement, and no agreement is better than an unfavorable agreement for each agent. This also means that both agents must be sufficiently optimistic that they are favored by the agreement in order to participate. Too much pessimism will cause either agent to opt out, thus learning too much about the agreement can be detrimental.

Agents share a common prior belief p_0 that agent 1 is favored, and a common posterior belief p_t at each instant t. Given each player's actions $\{k_{i,t}\}_{t=0}^{\infty}$ such that $k_{i,t}$ is measurable with respect to the information available at time t, player 1's total expected discounted payoff, expressed in per-period units, is

$$\mathbb{E}\left[\int_0^\infty re^{-rt}[K_t\lambda_1 p_t h + (1 - K_t)s]dt\right] ,$$

and player 2's total expected discounted payoff, expressed in per-period units, is

$$\mathbb{E}\left[\int_0^\infty re^{-rt}[K_t\lambda_2(1 - p_t)h + (1 - K_t)s]dt\right] ,$$

where $K_t = k_{1,t}k_{2,t}$ indicates if the agreement is in force or not.

[3] We can also let the lump-sum reward to be uncertain, with h on average. This would not change our results.

Learning There are three possible events that may occur over time interval dt. First, a reward arrives only for agent 1, so that $p_{t+dt} = 1$. Second, a reward arrives only for agent 2, so that $p_{t+dt} = 0$. Lastly, no reward arrives for either agent, so that beliefs are updated via Bayes' rule according to[4]

$$p_{t+dt} = \frac{p_t e^{-K_t \lambda_1 dt}}{p_t e^{-K_t \lambda_1 dt} + (1 - p_t) e^{-K_t \lambda_2 dt}} .$$

The novelty of our approach lies in K_t which is zero whenever one of the agents chooses to discontinue the agreement, i.e. $k_{i,t} = 0$ for some i. The implication is that when one agent discontinues the agreement, learning stops for both agents. In other words, agent 1 needs agent 2's continued participation to learn, and vice versa. Therefore, our approach is not centered around free-riding problems, but rather around a common incentive to learn that slowly disappears as information is acquired.

Equilibrium We restrict attention to pure Markov strategies with the belief $p \in [0, 1]$ as the payoff relevant state. We denote a Markov strategy for agent i as κ_i, which is a choice to remain in the agreement or not conditional on belief p:

$$\kappa_i : [0, 1] \to \{0, 1\} .$$

As in Klein and Rady (2011), each strategy pair (κ_1, κ_2) induces a pair of value functions

$$u_1(p|\kappa_1, \kappa_2) = \mathbb{E}\left[\int_0^\infty re^{-rt}[\kappa_1(p_t)\kappa_2(p_t)\lambda_1 p_t h + (1 - \kappa_1(p_t)\kappa_2(p_t))s]dt \middle| p_0 = p\right]$$

and

$$u_2(p|\kappa_1, \kappa_2) = \mathbb{E}\left[\int_0^\infty re^{-rt}[\kappa_1(p_t)\kappa_2(p_t)\lambda_2(1 - p_t)h + (1 - \kappa_1(p_t)\kappa_2(p_t))s]dt \middle| p_0 = p\right]$$

for players 1 and 2, respectively. A Markov perfect equilibrium is a pair of strategies (κ_1, κ_2) such that for each agent i, $\kappa_i(p)$ maximizes her value function for all p given κ_j. In order to eliminate trivial equilibria in which neither agent participates in the agreement, we assume agents break indifference in favor of participating in the agreement.[5] We henceforth refer to a Markov equilibrium simply as an equilibrium.

Equilibrium in cut-off strategies Consider an equilibrium such that agent i participates in the agreement as long as she is sufficiently optimistic. Define cutoffs $\overline{p}_1$ and $\overline{p}_2$ such

[4] The probability of a reward arriving for both agents or multiple rewards arriving for the same agent is negligible as the terms of order $o(dt)$ can be ignored.

[5] Standard refinements, such as payoff dominance can also eliminate such equilibria without the need to assume a tie-breaking rule.

that agent 1 remains in the agreement if $p \geq \overline{p}_1$ and agent 2 remains in the agreement if $p \leq \overline{p}_2$. Therefore, the agreement is only in force if p lies in the experimentation region $[\overline{p}_1, \overline{p}_2]$. By standard arguments, while the agreement is in force, agents' value functions satisfy the following ordinary differential equations (ODEs):

$$\begin{aligned} ru_1(p) &= r\lambda_1 ph - (\lambda_1 - \lambda_2)p(1-p)u_1'(p) \\ &\quad + \lambda_1 p(s - u_1(p)) + \lambda_2(1-p)(s - u_1(p)) \ , \end{aligned} \tag{1}$$

$$\begin{aligned} ru_2(p) &= r\lambda_2(1-p)h - (\lambda_1 - \lambda_2)p(1-p)u_2'(p) \\ &\quad + \lambda_1 p(s - u_2(p)) + \lambda_2(1-p)(s - u_2(p)) \ . \end{aligned} \tag{2}$$

Notice that for agent 1, a reward arrives with probability proportional to $\lambda_1 p$, resulting in an instantaneous payoff of rh, however, agent 1 pays opportunity cost $u_1(p) - s$ since the agreement dissolves immediately. A reward arrives for agent 2 with probability proportional to $\lambda_2(1-p)$. In that case, agent 1 only pays opportunity cost $u_1(p) - s$. Additionally, there is a downward drift resulting from the evolution of p in the absence of any reward, captured by the term $-(\lambda_1 - \lambda_2)p(1-p)u_1'(p)$. While the agreement is in force and no news arrive, the posterior belief solves the ODE $dp = -p(1-p)\Delta_\lambda dt$, where $\Delta_\lambda = \lambda_1 - \lambda_2$ captures the asymmetry of judgment frequencies in the agreement. When $\Delta_\lambda > 0$, no arrival of rewards is evidence against agent 1 having the good risky arm and results in a downward drift in beliefs. In a more informative agreement, no arrival of rewards is stronger evidence against agent 1 and thus makes belief drift faster, i.e., the speed of learning increases. Since beliefs drift down while the agreement is in force, agent 1 becomes more pessimistic that he is favored as time goes on, but agent 2 becomes more optimistic. For beliefs sufficiently low, agent 1 prefers to exit the agreement and take the safe payoff, while for beliefs sufficiently high agent 2 prefers not to begin the agreement in the first place. The threshold belief $\overline{p}_1$ is such that agent 1 is indifferent between continuing the agreement one more instant, and exiting the agreement, while agent 2's threshold belief $\overline{p}_2$ is such that agent 2 is indifferent between remaining with no agreement, or *starting* the agreement in the first place. Thus at $\overline{p}_i$ agent i's value function is equal to the dynamic payoff from the safe option for each i.

Myopic threshold beliefs As a natural benchmark, consider agents' myopic cut-off beliefs. These are such that the agent is indifferent between the instantaneous payoff to be being outside the agreement s, and the expected instantaneous payoff to being in the agreement, $\lambda_1 ph$, for agent 1 and $\lambda_2(1-p)h$ for agent 2. The myopic thresholds are

$$p_1^m = \frac{s}{\lambda_1 h} \qquad \text{and} \qquad p_2^m = 1 - \frac{s}{\lambda_2 h} \ .$$

Note that $\lambda_1 \geq \lambda_2$ and $\lambda_2 h > 2s$ imply $p_1^m < p_2^m$, so that there exists some p for which the agreement would be formed by myopic agents.

As can be expected, the myopic benchmark for agent 1 increases with the safe payoff s and decreases with player 1's expected payoff conditional on being favored, $\lambda_1 h$. Similarly, agent 2's myopic benchmark decreases with s and increases with player 2's expected payoff conditional on being favored, $\lambda_2 h$. As the safe payoff increases, the interval of beliefs such that the agreement is in force decreases because the incentive to experiment with the agreement is lower. The reverse is true for h as the expected payoff to being in the agreement, conditional on being favored, increases.

While the interval of beliefs for myopic players increases with λ_i, it is not clear that in a dynamic setting the agreement will last longer as λ_i increases. From the social planner's perspective, a longer lasting agreement is always better, so the longevity of the agreement is of concern. In a dynamic setting, this will not only depend on the incentive to exit the agreement, but also the speed of learning and the prior belief.

Agreement value For an equilibrium in cut-off strategies, we define the *agreement value* $V(p_0)$ as the expected joint surplus of agent 1 and agent 2, conditional on the initial belief p_0. If the agents initiate the agreement at the initial belief p_0, that is, $\overline{p}_1 \leq p_0 \leq \overline{p}_2$, then the agreement value is given by the sum of their equilibrium value functions, $V(p_0) = u_1(p_0) + u_2(p_0)$. Otherwise, the agreement is never started, and the agreement value is equal to a sum of safe payoffs, $V(p_0) = 2s$. The agreement value captures the private component of welfare: it measures the total surplus generated by the two agents when both find it optimal to participate, taking into account their strategic exit incentives and the evolution of beliefs.

Societal value Assume that the agreement, while in force, exerts a flow externality b for society. We study an economically interesting and relevant case of a positive externality, $b > 0$. For an equilibrium in cut-off strategies, we define the *societal value* $S(p_0)$ as the expected present discounted value of the flow externalities (expressed in per-period units) generated by the agreement, conditional on the initial belief p_0. If the agents initiate the agreement at the initial belief p_0, that is, $\overline{p}_1 \leq p_0 \leq \overline{p}_2$, then the agreement value is

$$S(p_0) = \mathbb{E}\left[\int_0^\tau rbe^{-rt}dt\right] ,$$

where τ denotes the random stopping time at which the agreement terminates. If the agreement is never formed, the societal value is $S(p_0) = 0$, as no externalities are ever generated. The societal value captures the social component of welfare: it measures the total discounted benefit to parties outside the agreement while it remains in force.

Welfare We define welfare $W(p)$ as a linear combination of the agreement value and the societal value, $W(p) = (1-\beta)V(p) + \beta S(p)$, where $\beta \in [0,1]$ captures the relative weight placed on the societal value in the welfare objective.

We study two dimensions of welfare maximization. The *welfare-maximizing prior* p^* corresponds to the belief at which the agreement is most valuable ex ante, given fixed judgment frequencies λ_1, λ_2. The *welfare-maximizing judgment frequencies*, λ_1^*, λ_2^* characterize the pair of Poisson process intensities that maximize welfare for a given prior p_0. When $\lambda_i^* = +\infty$ is optimal, we say that instant judgments are welfare-maximizing. These objects capture how a potential agreement designer would ideally structure the environment to balance incentives to participate and the informational efficiency of the agreement.

In practice, choosing the prior can be interpreted as deciding when negotiations should crystallize into an agreement. That is, how much evidence or consensus in favor of one side should be accumulated before formal cooperation starts.[6] Choosing judgment frequencies corresponds to shaping the institutional or procedural features that govern how frequently decisions or rulings occur. In the context of international or regulatory agreements, judgment frequencies reflect the design of dispute-resolution mechanisms, or the rate at which cases are adjudicated. While we do not explicitly model the designer as a strategic player, studying welfare-maximizing priors and judgment frequencies reveals how institutional design features affect the overall efficiency and stability of the agreement.

4 Symmetric Agents

To develop intuition, we begin with the case of symmetric judgment frequencies in which the distribution of rewards conditional on p is identical for both agents, i.e., $\lambda_1 = \lambda_2 = \lambda$. With equal arrival rates, the absence of rewards conveys no new information, and beliefs remain constant: $dp = 0$. Intuitively, from the perspective of agent i, no payoff arriving for her is bad news, while no payoff arriving for agent j is good news of the exact same magnitude. The informativeness and the speed of learning Δ_λ in the absence of rewards are zero. Simplifying agents' value functions (1) and (2), we obtain the following linear equations:

$$\begin{aligned} ru_1(p) &= r\lambda ph + \lambda p(s - u_1(p)) + \lambda(1-p)(s - u_1(p)) \, , \\ ru_2(p) &= r\lambda(1-p)h + \lambda p(s - u_2(p)) + \lambda(1-p)(s - u_2(p)) \, . \end{aligned}$$

[6]Understanding the welfare-maximizing priors can also be useful from the information design perspective, where the designer may generate public signals to move agents' starting belief closer to the welfare-maximizing one.

These equations have straightforward solutions for agents' value functions:

$$u_1(p) = \frac{\lambda(prh + s)}{\lambda + r} ,$$
$$u_2(p) = \frac{\lambda((1-p)rh + s)}{\lambda + r} .$$

Note that $u_1(p)$ is strictly increasing in p, while $u_2(p)$ is strictly decreasing in p as each agent's value from being in the agreement decreases with their pessimism about being the favored agent. Agents' cutoff beliefs satisfy $u_1(\overline{p}_1) = s$ and $u_2(\overline{p}_2) = s$. Value matching must hold since at the cutoff belief an agent must be indifferent between continuing the agreement and exiting for the safe payoff. These simple conditions yield that both agents' equilibrium cutoffs coincide with their myopic cutoffs. We summarize this result in the following proposition.

Proposition 1 *In the case of symmetric agents, there is a unique cutoff-equilibrium in which the agreement is in force if and only if beliefs lie between the agent's myopic thresholds, i.e., if* $\overline{p}_1 \leq p \leq \overline{p}_2$*, where* $\overline{p}_1 = p_1^m$ *and* $\overline{p}_2 = p_2^m$ *.*

Agents place no value on further experimentation for beliefs less optimistic than p_i^m, as no learning under no news occurs and breakthroughs resolve uncertainty. The equilibrium in Proposition 1 describes an agreement that forms if and only if myopic agents would find the agreement worthwhile, that is $p_1^m \leq p_0 \leq p_2^m$, and ends when there is a breakthrough for either player. Such a breakthrough occurs with probability λdt over a small time interval dt. Since p_1^m is decreasing in λ while p_2^m is increasing in λ, a higher λ expands the set of priors for which the agreement can be sustained.

We characterized the equilibrium under symmetric judgment frequencies. We now turn to the welfare properties, beginning with the agreement and societal values.

The agreement value depends on the initial belief. If the agents initiate the agreement at the initial belief p_0, that is, $p_1^m \leq p_0 \leq p_2^m$, then the agreement value is given by the sum of agents' value functions, $u_1(p_0) + u_2(p_0)$. Otherwise, the agreement is never formed, and the agreement value is equal to the sum of safe payoffs, $2s$. Thus,

$$V(p_0) = \begin{cases} \frac{\lambda}{r+\lambda}(rh + 2s) , & \text{if } p_1^m \leq p_0 \leq p_2^m , \\ 2s , & \text{otherwise} . \end{cases}$$

If p_0 lies in the experimentation region $[p_1^m, p_2^m]$, then $\lim_{\lambda \to \frac{2s}{h}} V(p_0) = 2s$ and $\lim_{\lambda \to \infty} V(p_0) = rh + 2s$. Define $\overline{\lambda} = \max\left\{\frac{s}{p_0 h}, \frac{s}{(1-p_0)h}\right\}$ as the cutoff on λ, such that if $\lambda < \overline{\lambda}$, one of the players prefer to exit at p_0. Equivalently, $\overline{\lambda}$ shows the minimal symmetric judgment frequency under which the agreement is started at the initial belief p_0. Figure 1 shows the agreement value for fixed prior belief p_0 across different λ. Given that λ is high enough so that the agreement starts, the agreement value increases in λ, as the instant judgments are beneficial to quickly resolve uncertainty.

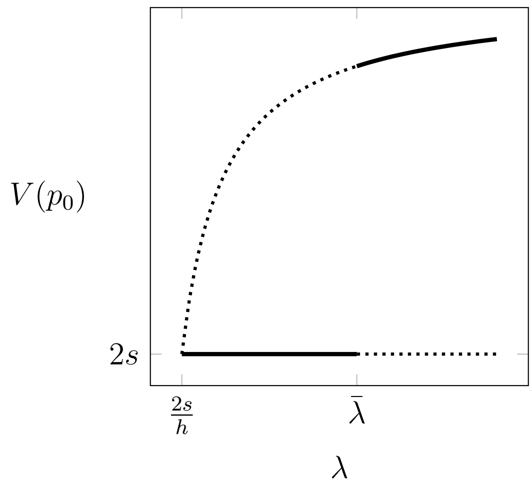

Figure 1: Agreement value $V(p_0)$ for fixed p_0 as a function of λ.

The societal value also depends on whether the starting belief p_0 lies in the experimentation region. If the agents initiate the agreement at p_0, that is, $p_1^m \le p_0 \le p_2^m$, then the societal value is given by the expected present-discounted value of the flow externalities until the first breakthrough. Thus, given that the agreement is in force at p_0, the societal value is

$$S(p_0) = \mathbb{E}\left[\int_0^\tau rbe^{-rt}dt\right] = \frac{r}{r+\lambda}b \ ,$$

where τ is the stopping time of the first breakthrough. Otherwise, the agreement is never formed, no externalities arrive, and $S(p_0) = 0$. Thus,

$$S(p_0) = \begin{cases} \frac{r}{r+\lambda}b \ , & \text{if } p_1^m \le p_0 \le p_2^m \ , \\ 0 \ , & \text{otherwise} \ . \end{cases}$$

For values of λ below $\overline{\lambda}$, the agreement never starts and does not realize any societal value. As long as the agreement starts, the societal value decreases in λ as the agreement stays in place shorter, with $\lim_{\lambda\to\infty} S(p_0) = 0$ in the limit. Figure 2 shows the societal value for fixed prior belief p_0 across different λ.

Welfare is a weighted average of the private and social components of value, $W(p_0) = (1-\beta)V(p_0) + \beta S(p_0)$. Combining our observations for the agreement and societal values, we have an explicit expression for welfare:

$$W(p_0) = \begin{cases} (1-\beta)\frac{\lambda}{r+\lambda}(rh+2s) + \beta\frac{r}{r+\lambda}b \ , & \text{if } p_1^m \le p_0 \le p_2^m \ , \\ (1-\beta)2s \ , & \text{otherwise} \ . \end{cases}$$

We can immediately conclude that a welfare-maximizing prior p^* under fixed symmetric judgment frequency λ is any p^* lying in the experimentation region $[p_1^m, p_2^m]$. Indeed, the welfare function is flat over this interval: since agents are symmetric, all such prior beliefs

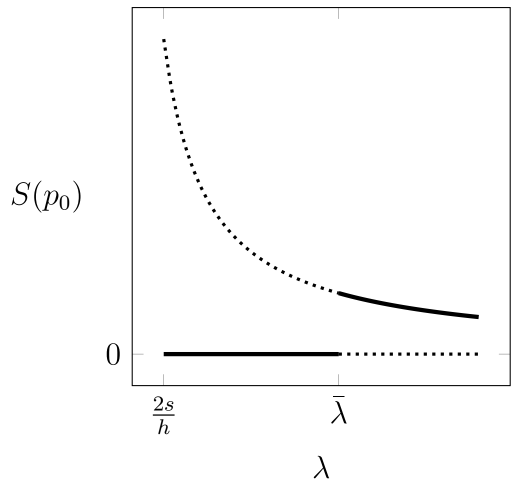

Figure 2: Societal value $S(p_0)$ for fixed p_0 as a function of λ.

lead to identical outcomes, with agreement formation followed by termination upon the first breakthrough.

We now turn to the welfare-maximizing symmetric judgment frequency λ^*, fixing prior belief p_0. The welfare expression makes transparent the fundamental trade-off between the speed of resolution and the longevity of the agreement. A higher judgment frequency increases the agreement value by accelerating resolution but reduces its societal value by shortening its expected duration. For $p_1^m \leq p_0 \leq p_2^m$, we can rewrite the welfare function isolating the effect of symmetric judgment frequency:

$$W(p_0) = (1-\beta)(rh+2s) - [(1-\beta)(rh+2s) - \beta b] \cdot \frac{r}{r+\lambda} \; .$$

The solution to the maximization problem $W(p_0) \to \max_\lambda$ is then immediate and depends on the sign of $(1-\beta)(rh+2s) - \beta b$ term:

$$\lambda^* = \begin{cases} +\infty \; , & \text{if } (1-\beta)(rh+2s) - \beta b \geq 0 \; , \\ \bar{\lambda} \; , & \text{if } (1-\beta)(rh+2s) - \beta b \leq 0 \; . \end{cases}$$

We summarize the result in Propositions below beginning with the case when the societal value has a low weight in total welfare.

Proposition 2 *Suppose societal value has a relatively low weight in total welfare, i.e.,* $\beta \leq \frac{rh+2s}{b+rh+2s}$. *Then*

1. *Instant judgments are optimal.*

2. *When judgments are instant, the value function for welfare is* $W^*(p_0) = (1-\beta)(rh+2s)$.

3. *With instant judgments, an agreement is always in place and instantly resolved, i.e.,* $\lambda \to \infty$, $p_1^m \to 0$ *and* $p_2^m \to 1$.

The case of the high societal value is quite different, and depends on the agents' prior belief.

Proposition 3 *Suppose societal value has a relatively high weight in total welfare, i.e.,* $\beta \geq \frac{rh+2s}{b+rh+2s}$*. Then*

1. *Optimal judgments are as slow as possible while maintaining agents' incentives to initiate the agreement, i.e.,* $\lambda^* = \overline{\lambda}(p_0) = \max\left\{\frac{s}{p_0 h}, \frac{s}{(1-p_0)h}\right\}$.
2. *With the optimal judgment frequency, the value function for welfare is* $W^*(p_0) = (1-\beta)(rh+2s) - [(1-\beta)(rh+2s) - \beta b] \cdot \frac{r}{r+\overline{\lambda}(p_0)}$*, which is symmetric around* $p_0 = 1/2$.
3. *An agreement is always formed and the judgments arrive with frequency* $\overline{\lambda}(p_0)$.

When welfare places little weight on the societal value, instant judgments are optimal because they resolve uncertainty immediately and maximize the agreement value. When the societal externality is sufficiently important, slow and deliberate judgments are preferred: prolonged agreement sustains cooperation and generates greater societal value. This trade-off between speed and longevity of agreements provides a benchmark for understanding the role of asymmetry in the next section.

5 Asymmetric Agents

We now consider the central case of asymmetric agents. To understand the role of resulting belief drift, we characterize the equilibrium cutoff beliefs and value functions when agents differ in judgment frequencies, i.e., $\lambda_1 > \lambda_2$. Since $\Delta_\lambda > 0$, the agreement becomes informative and beliefs drift down while the agreement is in force and no rewards arrive for any agent. Fixing equilibrium cutoffs, Lemma 1 characterizes the agents' value functions up to constants of integration.

Lemma 1 *In the experimentation region* $p \in [\overline{p}_1, \overline{p}_2]$*, value functions take the following forms:*

$$\begin{aligned} u_1(p) &= C_1(1-p)(\Omega(p))^\alpha + A_1 p + B_1(1-p) \\ u_2(p) &= C_2(1-p)(\Omega(p))^\alpha + A_2 p + B_2(1-p) \end{aligned}$$

where $\Omega(p) = \frac{1-p}{p}$ *is the odds ratio at belief* p, $\alpha = \frac{r+\lambda_2}{\lambda_1-\lambda_2}$,

$$A_1 = \frac{\lambda_1(rh+s)}{r+\lambda_1}, \qquad B_1 = \frac{\lambda_2 s}{r+\lambda_2}$$

for agent 1,

$$A_2 = \frac{\lambda_1 s}{r+\lambda_1}, \qquad B_2 = \frac{\lambda_2(rh+s)}{r+\lambda_2}$$

for agent 2, and C_1 *and* C_2 *are constants of integration to be determined.*

The non-linear term $C_i(1-p)(\Omega(p))^\alpha$ captures the option value associated with being able to exit the agreement, while the linear term $A_i p + B_i(1-p)$ represents the expected payoff from committing to the agreement given belief p. In the symmetric case there is no option value of being able to exit, as the beliefs do not move when there is no news. When $\lambda_1 > \lambda_2$, this is no longer a case. Intuitively, we have $A_1 > s > B_1$, as the agent 1's expected payoff from committing to the agreement given that agent 1 is favored exceeds the safe payoff s, while the expected payoff from committing given that agent 2 is favored is below s. Similarly, $A_2 < s < B_2$ for agent 2. The linear component of the value function when $\lambda_1 = \lambda_2$ coincides with the value function in the symmetric case, since the expected payoff from committing to the agreement is not affected by belief drift.[7]

For agent 1, under the downward belief drift, standard value matching and smooth pasting conditions pin down agent 1's cutoff belief $\overline{p}_1$ and the constant of integration C_1. These conditions are $u_1(\overline{p}_1) = s$ and $u_1'(\overline{p}_1) = 0$, respectively. The following proposition provides the equilibrium solution for agent 1.

Proposition 4 *When $\lambda_1 > \lambda_2$, the agent 1's threshold belief coincides with his myopic threshold, i.e., $\overline{p}_1 = p_1^m$, and the constant of integration C_1 is positive and solves*

$$C_1 \left(\frac{\lambda_1 h - s}{s}\right)^\alpha = \frac{r(\lambda_1 - \lambda_2)s}{(r+\lambda_1)(r+\lambda_2)} .$$

To see why the agent 1's equilibrium cutoff is the same as the myopic cutoff, note that the arrival of a reward for agent 1 immediately terminates the agreement. A reward for agent 1 is conclusive news that agent 2 is not favored, hence agent 2 no longer has an incentive to participate. For agent 1, there is no subsequent value of experimentation as a result, and only myopic incentives are relevant, as belief drift is unfavorable to agent 1. This will not be the case for agent 2 due to the direction of belief drift. Observe that agent 1's constant of integration C_1 is positive and value function u_1 is convex and achieves a minimum at $\overline{p}_1$ on the experimentation region. Agent 1's value is thus eroded over time while the agreement is in force. Lemma 3 in Appendix A shows that $\lim_{\Delta_\lambda \downarrow 0} C_1 = 0$, confirming our intuition that the option value goes to zero when asymmetry dissipates.

We now turn to agent 2. Because beliefs drift downward in the absence of breakthroughs, agent 2 faces a more favorable informational environment: under no news, the posterior gradually moves in her favor. Thus, for a given prior p_0, agent 2 either never enters the agreement or remains willing to participate until the first reward arrives. What remains is to determine the constant of integration C_2 and the corresponding cutoff belief $\overline{p}_2$.

Unlike agent 1's, agent 2's value function no longer satisfies the continuity of smoothness at $\overline{p}_2$. The reason is that the boundary of the experimentation region $\overline{p}_2$ is not regular: if the agreement starts at $p_0 = \overline{p}_2$, the posterior belief does not enter the stopping region immediately

[7] For example, we can rewrite agent 1's value function as $u_1(p) = \frac{\lambda_1(rh+s)}{r+\lambda_1}p + \frac{\lambda_2 s}{r+\lambda_2}(1-p)$ in the symmetric case.

with probability 1.[8] Instead, for agent 2 to participate, the value of the agreement must be at least as great as the value from staying out. Thus the threshold belief must satisfy the continuous pasting condition, $u_2(\overline{p}_2) = s$. However, due to downward belief drift it must also be true that when agent 1 decides to quit the agreement, agent 2's value falls to s, $u_2(\overline{p}_1) = s$. The latter equilibrium condition pins down agent 2's constant of integration.

Lemma 2 *The constant of integration for agent 2's value function is negative and solves*

$$C_2 \;=\; -\frac{C_1}{\lambda_1 h - s}\left[\frac{(r+\lambda_1)h[\lambda_1\lambda_2 h - s(\lambda_1+\lambda_2)]}{(\lambda_1-\lambda_2)s} + s\right] .$$

Lemma 2 completes the characterization of agent 2's value function. Moreover, we can see from Lemma 2 that C_2 is negative, hence u_2 is concave. The downward drift in beliefs gives agent 2 a positive option value as agent 2 expects the environment to improve the longer no breakthrough occurs. Consistent with the previous intuition, Lemma 3 in Appendix A shows that $\lim_{\Delta_\lambda \downarrow 0} C_2 = 0$. We illustrate agents' equilibrium value functions in Figure 3 along with the equilibrium cutoffs.

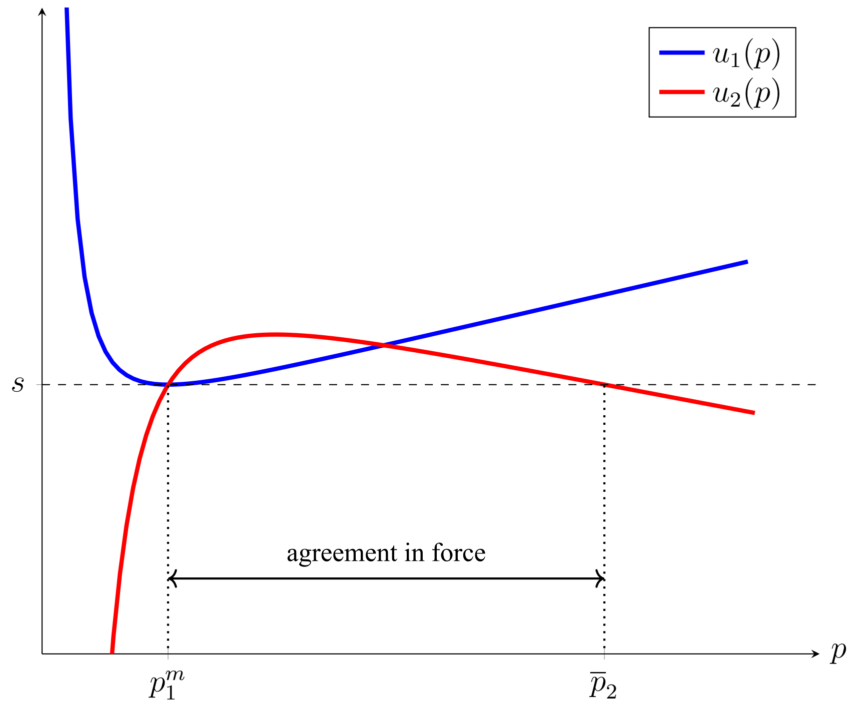

Figure 3: Agents' value functions and equilibrium cutoffs.

Intuitively, $u_2(p)$ is equal to safe payoff s at $\overline{p}_2$, and the value of C_2 is such that $u_2(p)$ also attains the safe payoff on the left at $\overline{p}_1 = p_1^m$. We formally characterize $\overline{p}_2$ in Proposition 5. We show that agent 2's equilibrium cutoff exceeds her myopic cutoff. We also show that $\overline{p}_2$ increases in judgment frequencies for mild asymmetries.

[8] See Peskir and Shiryaev (2006), Section 7, for the discussion of optimality conditions at regular and irregular boundaries. A similar failure of smooth pasting arises in the model of breakdowns in Keller and Rady (2015).

Proposition 5 *For* $\Delta_\lambda = \lambda_1 - \lambda_2 > 0$*,*

(i) agent 2 chooses to start the agreement for $p_0 \leq \overline{p}_2$*, where* $\overline{p}_2$ *is the maximum value that solves*

$$C_2(1-\overline{p}_2)(\Omega(\overline{p}_2))^\alpha + \frac{\lambda_1 s}{r+\lambda_1}\overline{p}_2 + \frac{\lambda_2(rh+s)}{r+\lambda_2}(1-\overline{p}_2) = s.$$

(ii) $\overline{p}_2 > p_2^m$ *and* $\overline{p}_2 \downarrow p_2^m$ *as* $\Delta_\lambda \downarrow 0$.

(iii) for Δ_λ *sufficiently small,* $\overline{p}_2$ *is increasing in* λ_1 *and* λ_2.

When the agents are symmetric, the agreement is dissolved only through an arrival of a reward for either agent 1 or 2. In the asymmetric case, now that there is a downward drift of beliefs, the agreement may be dissolved if for a sufficiently long time no rewards arrive due to lacking evidence favoring agent 1. Hence, fixing the total instantaneous probability of a breakthrough, $p\lambda_1 + (1-p)\lambda_2$, an asymmetric agreement is shorter lived than a symmetric one. Therefore, the opportunity cost, that is, foregoing a safe payoff of s for the duration of the agreement, is lower. This motivates player 2 to stay in the agreement beyond p_2^m. Moreover, when asymmetry vanishes, agent 2's dynamic incentives converge to the myopic ones.

Increasing λ_2 generates greater agent 2's payoff conditional on being favored, directly improving willingness to start the agreement. For small asymmetries, this direct payoff effect dominates any informational changes and agent 2 tolerates a larger probability that agent 1 is favored when deciding whether to start the agreement. Thus, agent 2's cutoff is increasing in λ_2. Increasing λ_1 increases Δ_λ and thus the speed of learning. This improves the value of experimentation for agent 2, since no news becomes increasingly good news. At the same time, increasing λ_1 also increases the rate at which rewards arrive for agent 1, which would terminate the agreement. Proposition 5 shows that the former effect dominates the latter for mild asymmetries, and agent 2's cutoff is increasing in λ_1.

The experimentation region therefore expands with asymmetry. An increase in λ_1 moves the equilibrium cutoffs away each other. The experimentation region also expands when λ_2 increases but only due to an upward shift in $\overline{p}_2$. Proposition 5 completes the characterization of the equilibrium in cut-off strategies. We now turn to the welfare properties.

As in the case of symmetric agents, the agreement value is the sum of agents' continuation values, $u_1(p_0) + u_2(p_0)$, if the agreement is in place at belief p_0. Otherwise, the agreement collapses, and both agents receive their safe payoffs. That is,

$$V(p_0) = \begin{cases} u_1(p_0) + u_2(p_0) \ , & \text{if } p_1^m \leq p_0 \leq \overline{p}_2 \ , \\ 2s \ , & \text{otherwise} \ . \end{cases}$$

When $p_1^m \leq p_0 \leq \overline{p}_2$, we have

$$V(p_0) = (C_1 + C_2)(1 - p_0)(\Omega(p_0))^{\alpha} + \left[\frac{\lambda_1}{r + \lambda_1} \cdot p_0 + \frac{\lambda_2}{r + \lambda_2} \cdot (1 - p_0)\right](rh + 2s) \ ,$$

where the constants of integration, $C_1 > 0$ and $C_2 < 0$, are given by Proposition 4 and Lemma 2, respectively. Importantly, Lemma 4 in Appendix A shows that their sum, $C_1 + C_2$, is negative whenever the agents are asymmetric. Although the downward drift in beliefs benefits agent 2 by improving her continuation value, it harms agent 1 more sharply by increasing the risk of premature termination of the agreement. Thus, asymmetry may not merely redistribute surplus between the agents but introduce a collective agreement value loss by making the agreement more fragile.

Given that the agreement is in force, the societal value solves the following ODE:

$$rS(p) = rb - \Delta_\lambda p(1 - p)S'(p) - \lambda_1 pS(p) - \lambda_2(1 - p)S(p) \ ,$$

with boundary condition $S(p_1^m) = 0$. Indeed, the instantaneous benefit for the societal value is rb, and any breakthrough stops the agreement. Solving this boundary-value problem, we obtain[9]

$$S(p_0) = \left(\frac{1 - p_0}{r + \lambda_2} + \frac{p_0}{r + \lambda_1}\right) rb - \left(\frac{1 - p_1^m}{r + \lambda_2} + \frac{p_1^m}{r + \lambda_1}\right) rb \left(\frac{1 - p_0}{1 - p_1^m}\right)^{\frac{\lambda_1 + r}{\Delta_\lambda}} \left(\frac{p_0}{p_1^m}\right)^{-\frac{\lambda_2 + r}{\Delta_\lambda}} \ ,$$

when $p_1^m \leq p_0 \leq \overline{p}_2$; and $S(p_0) = 0$, otherwise. Intuitively, the first linear term is what the society would get if the agreement ran at a fixed belief p_0, with discounting and stochastic Poisson termination. The second term is the dynamic correction due to the possibility that the agreement runs out before any breakthrough via exit by agent 1. As asymmetry increases, the belief drift accelerates and the likelihood of agreement termination increases. Thus, asymmetry may shorten the socially valuable phase of experimentation even when both agents remain willing to participate.

As before, welfare is a linear combination of the agreement value and the societal value:

$$W(p_0) = (1 - \beta)V(p_0) + \beta S(p_0) \ .$$

Asymmetry reshapes welfare by changing both the extent of the experimentation region and the informational dynamics within it. We are thus interested in the local and global effects of introducing asymmetry.

We proceed by characterizing the welfare-maximizing prior p^* fixing judgment frequencies λ_1 and λ_2. Our finding is organized around the *agreement stakes* $\frac{\lambda_i h}{s}$ defined as a measure of the payoff advantage of a good risky arm over a safe payoff. Note that we already assume that the

[9] See Appendix B for a detailed derivation of this expression.

agreements stakes are high enough so that myopic agents would start the agreement under some prior beliefs, $\frac{\lambda_i h}{s} > 2$.[10] We show that the agreement value is increasing and strictly concave over the experimentation region, while the societal value is also strictly concave but becomes hump-shaped if the agreement stakes are sufficiently high. In this case, the welfare-maximizing prior lies in the interior of the experimentation region whenever the welfare puts large enough weight on the societal value.

Proposition 6 *For fixed $\lambda_1 > \lambda_2$, there is a unique maximizer p^* of the welfare function W. For sufficiently high agreement stakes, S is hump-shaped, $S'(\overline{p}_2) < 0$, and there exists $\hat{\beta} \in (0, 1)$, such that $p^* = \overline{p}_2$ if $\beta \leq \hat{\beta}$ and $p^* \in (p_1^m, \overline{p}_2)$ if $\beta > \hat{\beta}$. If S is not hump-shaped, $S'(\overline{p}_2) \geq 0$, then $p^* = \overline{p}_2$ for any $\beta \in [0, 1]$.*

When the agreement stakes are high, the agreement generates substantial value both privately and socially.[11] In this case, the societal value rises and then falls over the experimentation region. Starting the agreement when beliefs are too pessimistic from agent 1's perspective reduces the agreement longevity, but starting when beliefs are too optimistic limits the duration of informative experimentation, since breakthroughs arrive quickly and resolve uncertainty. At low levels of concern for societal value (small β), the agreement value dominates, and welfare is maximized for the maximal belief at which both agents are willing to participate. When β is high, the designer values the longevity of experimentation more. The welfare-maximizing prior shifts into the interior of the experimentation region where the agreement is expected to last long enough to generate greater societal value.

When the agreement stakes are low, the negative effect of starting from high p_0 on the societal value weakens. In this regime, the societal value $S(p)$ increases monotonically with the prior belief, eliminating any benefit from starting with more pessimistic starting beliefs. As a result, welfare is maximized at the agent 2's cutoff $\overline{p}_2$ for all β. Appendix C provides an illustration of a strictly increasing societal value in the experimentation region in the environment with low agreement stakes.

Having characterized the welfare-maximizing prior for fixed judgment frequencies, we now turn to the second dimension of agreement design, judgment frequencies. Proposition 7 below establishes the key property of welfare-maximizing judgment frequencies λ_1^*, λ_2^* fixing prior belief p_0. For any asymmetric environment with judgment frequencies λ_1, λ_2, we can find a symmetric welfare-improving λ.

Proposition 7 *For fixed p_0, $\lambda_1 > \lambda_2$, there always exists λ, such that $W(p_0, \lambda_1, \lambda_2) \leq W(p_0, \lambda, \lambda)$. Thus, welfare-maximizing λ_1^*, λ_2^* are symmetric.*

[10] In a symmetric environment, this corresponds to the case of high stakes in Klein and Rady (2011).

[11] The proof of Proposition 6 provides an explicit lower bound on agreement stakes $\frac{\lambda_i h}{s} \gtrapprox 4.6$ to qualify a high-stake environment. This lower bound can be improved but it would depend on the parameters of the problem in a non-trivial way.

Asymmetry distorts how the benefits and risks of the agreement are shared. Because belief drift favors agent 2, agent 1's option value of exit rises, but this comes at the cost of a larger risk of premature breakdown from the perspective of society and agent 2. Proposition 7 shows that these option-value and drift effects sum to a loss: welfare is strictly lower under asymmetric $\lambda_1 > \lambda_2$ than under an appropriately chosen symmetric λ. In the symmetric case, there is no drift when no rewards arrive. Thus, beliefs remain constant until a breakthrough, and the agreement only ends when revealing information arrives in the form of a breakthrough. This keeps the agreement in place longer on average and lets both agents and society extract more value from the experimentation phase.

As a result, welfare-maximizing judgment frequences are symmetric, $\lambda_1^* = \lambda_2^* = \lambda^*$, and given by our analysis of symmetric agents in Proposition 2 and Proposition 3. In particular, when welfare places little weight on the societal value, instant judgments are optimal. When the societal value has a relatively high weight in welfare, optimal judgments are as slow as possible while maintaining agents' incentives to participate. Despite the global optimality of symmetric judgment frequencies, we show that introducing mild asymmetry locally can improve welfare.

To further outline the effect of asymmetry on welfare, starting from a fixed symmetric λ we introduce ρ-*asymmetry* defined by a strictly increasing function $\rho(\mu)$, such that if $\lambda_1 = \lambda + \mu$ and $\lambda_2 = \lambda - \rho(\mu)$, then $\bar{p}_2 = 1 - \frac{s}{\lambda h}$. That is, an increase in λ_1 by μ combined with and a corresponding decrease in λ_2 by $\rho(\mu)$ preserves agent 2's incentives to initiate the agreement, keeping agent 2's equilibrium cutoff at the level of the symmetric myopic cutoff. Function ρ is well defined, unique, and increasing for sufficiently small $\mu \geq 0$ by Proposition 5.

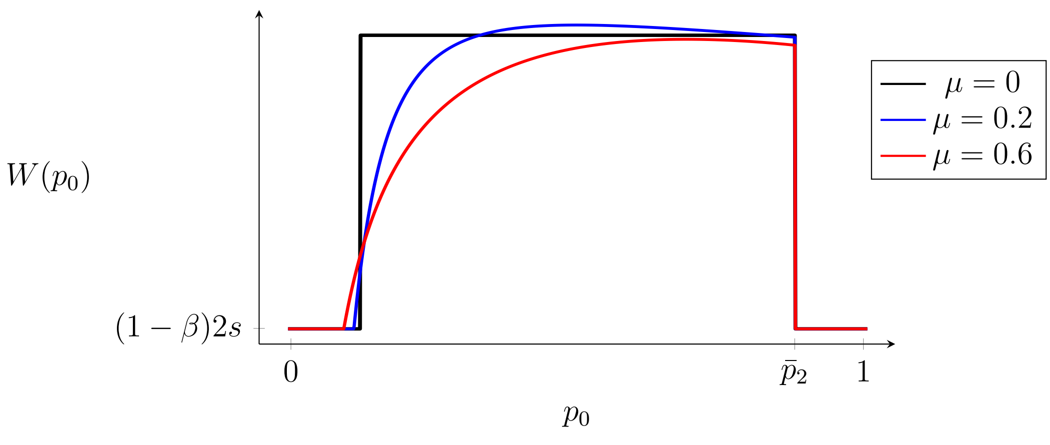

Figure 4: Welfare function for $\lambda_1 = \lambda + \mu$, $\lambda_2 = \lambda - \rho(\mu)$, $\lambda = 1, \beta = 0.5, b = 3, s = 0.5, h = 2, r = 0.1$.

A mild introduction of asymmetry via ρ-asymmetry increases welfare for a range of prior beliefs. Figure 4 showcases this result by comparing three welfare functions corresponding to: no asymmetry, mild asymmetry, and larger asymmetry. We can distinguish two ranges of prior beliefs where introducing asymmetry is welfare-improving. First, ρ-asymmetry improves welfare for p_0 to the left of agent 1's symmetric cutoff by strengthening agent 1's incentives to participate. Second, within the interior of the experimentation region under symmetry, mild asymmetry improves welfare by making the agreement more informative and dynamically valuable. A small drift in beliefs creates gradual updating even in the absence of breakthroughs, allowing the agents to learn while no breakthrough arrive. In this range, the welfare function becomes hump-shaped, and mild asymmetry enhances the informational efficiency of the agreement, while stronger asymmetry shortens its expected duration and eventually reduces welfare.

To summarize, although globally asymmetry lowers welfare by making agreements more fragile, locally it can improve welfare through two mechanisms: (1) by expanding the range of priors supporting participation and (2) by increasing dynamic informational value of the agreement when asymmetry is mild. This distinction is particularly relevant in environments where the agreement design permits adjusting only one side of judgment intensity or faces technological constraints on equalizing judgment frequencies.

6 Concluding Remarks

We use strategic experimentation to model the dynamics of risky agreements. The novelty of our approach is that agreement payouts require coordination: if one agent stops experimenting, i.e., quits the agreement, experimentation stops for all agents. Our focus is on the role of beliefs in the formation, duration and dissolution of agreements. We show that agreement formation generally requires beliefs to be moderate. Too extreme beliefs prohibit participation of pessimistic agents. The longevity of an agreement depends on how beliefs evolve, which in turn is governed by the frequency and informativeness of judgments. Agreements end when beliefs drift too far in one direction, triggering an agent's exit. In this environment, welfare accounts for the agents' value and a positive externality the agreement generates for actors outside the agreement. We show that agreements are generally inefficient. Many agreements, although preferred by a utilitarian social planner's perspective, will not form because of too extreme beliefs or be dissolved too early. This has important implications for the design and duration of agreements.

In the case of symmetric judgment frequencies, welfare involves a fundamental trade-off between the speed of resolution and the longevity of cooperation: rapid judgments favor private value, while slow judgments enhance societal value. Optimal design therefore depends on how much weight welfare places on the external benefits of continued agents' cooperation. When agents are asymmetric, belief drift introduces learning and dynamic incentives in the agreement. While global welfare is maximized by symmetric judgment frequencies, mild asymmetry can locally enhance welfare. These insights suggest that imperfect symmetry in institutional design, such as slightly unbalanced judgment or dispute-resolution frequencies, can promote participation and agreement stability. Our results highlight that the fragility and welfare performance of risky agreements depend critically on how institutional parameters shape the evolution of beliefs. Designing agreements that balance informational efficiency with longevity is key to sustaining cooperative arrangements in risky environments.

References

Bernheim, B. D. and Whinston, M. D. (1998). Incomplete contracts and strategic ambiguity. *The American Economic Review*, 88(4):902–932.

Bolton, P. and Harris, C. (1999). Strategic experimentation. *Econometrica*, 67(2):349–374.

CBC (2023). In arguably biggest test of new nafta, canada and mexico defeat u.s. in auto rules dispute. *CBC News Online*. Available at: https://www.cbc.ca/news/politics/cusma-rules-origin-autos-case-1.6710190 (Accessed: March 21, 2024).

Chiappori, P. and Salanie, B. (2013). *Asymmetric Information in Insurance Markets: Predictions and Tests*, pages 397–422.

Dütting, P., Feldman, M., Peretz, D., and Samuelson, L. (2024). Ambiguous contracts. *Econometrica*, 92(6):1967–1992.

Eliaz, K. and Spiegler, R. (2006). Contracting with diversely naive agents. *The Review of Economic Studies*, 73(3):689–714.

Giat, Y. and Subramanian, A. (2013). Dynamic contracting under imperfect public information and asymmetric beliefs. *Journal of Economic Dynamics and Control*, 37(12):2833–2861.

Grant, S., Kline, J. J., and Quiggin, J. (2014). A matter of interpretation: Ambiguous contracts and liquidated damages. *Games and Economic Behavior*, 85(C):180–187.

Grimmelmann, J. (2019). All Smart Contracts Are Ambiguous. *Journal of Law and Innovation, Cornell Legal Studies Research Paper No. 19-20,*, 1(1).

Hörner, J., Klein, N., and Rady, S. (2022). Overcoming free-riding in bandit games. *The Review of Economic Studies*, 89(4):1948–1992.

Keller, G. and Rady, S. (2010). Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311.

Keller, G. and Rady, S. (2015). Breakdowns. *Theoretical Economics*, 10(1):175–202.

Keller, G., Rady, S., and Cripps, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68.

Klein, N. and Rady, S. (2011). Negatively correlated bandits. *The Review of Economic Studies*, 78(2):693–732.

Peskir, G. and Shiryaev, A. (2006). *Optimal stopping and free-boundary problems*. Springer.

Riedel, F. and Sass, L. (2013). The strategic use of ambiguity. Working paper.

Riedel, F. and Sass, L. (2014). Ellsberg games. *Theory and Decision*, 76:469–509.

Strulovici, B. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica*, 78(3):933–971.

Tillio, A. d., Kos, N., and Messner, M. (2016). The Design of Ambiguous Mechanisms. *The Review of Economic Studies*, 84(1):237–276.

USTR (2020). United states-mexico-canada agreement. *Online*. Available at: https://ustr.gov/trade-agreements/free-trade-agreements/united-states-mexico-canada-agreement (Accessed: March 21, 2024).

Appendix A: Proofs

Proof of ***Proposition 1**, **Proposition 2**, **Proposition 3**:*

Proof. The proofs follow directly from the arguments provided in the main text. ■

Proof of ***Lemma 1**:*

Proof. Given the learning process and the shape of ODEs for value functions, equations (1) and (2), the general shape of the agent i's value function is a standard guess in strategic experimentation models with exponential bandits: see, e.g., Keller et al. (2005) and Hörner et al. (2022) for more general processes. The space of homogeneous solutions of ODEs is one-dimensional, and the general solution is the sum of an affine particular solution and a nonlinear term of the form $C_i(1-p)(\Omega(p))^\alpha$. As a result, the exponent α is the same for both agents, since it solves the homogeneous ODE. Given agent i's value function, $u_i(p) = C_i(1-p)(\Omega(p))^\alpha + A_i p + B_i(1-p)$, we have the following expression for the marginal value:

$$u_i'(p) = -C_i\left(1+\frac{\alpha}{p}\right)(\Omega(p))^\alpha + A_i - B_i\ .$$

For agent 1, substituting $u_1(p)$ and $u_1'(p)$ in equation (1) and matching the coefficients results in

$$(r+\lambda_1 p+\lambda_2(1-p))C_1(1-p) = \Delta_\lambda p(1-p)C_1\frac{p+\alpha}{p}$$

for the coefficient on $(\Omega(p))^\alpha$ and

$$(r+\lambda_2)A_1 p + \Delta_\lambda A_1 p + (r+\lambda_2)B_1(1-p) = r\lambda_1 hp + \lambda_1 sp + \lambda_2 s(1-p)$$

for the remaining terms. The first equation leads to $\alpha = \frac{r+\lambda_2}{\Delta_\lambda}$, while matching the coefficients for p- and $(1-p)$-terms in the second equation leads to $A_1 = \frac{\lambda_1(rh+s)}{r+\lambda_1}$ and $B_1 = \frac{\lambda_2 s}{r+\lambda_2}$. The derivation for agent 2 follows similar steps and thus omitted. ■

Proof of ***Proposition 4**:*

Proof. Agent 1's cutoff $\overline{p}_1$ and constant of integration C_1 are pinned down by the value-matching condition $u_1(\overline{p}_1) = s$ and the smooth-pasting condition $u_1'(\overline{p}_1) = 0$. From the smooth-pasting condition, we have the following expression for the option-value term of the value function:

$$C_1(1-\overline{p}_1)(\Omega(\overline{p}_1))^\alpha = \frac{\overline{p}_1(1-\overline{p}_1)}{\overline{p}_1+\alpha}(A_1-B_1)\ .$$

Combining this observation with the value-matching condition, we obtain

$$\frac{\overline{p}_1(1-\overline{p}_1)}{\overline{p}_1+\alpha}(A_1-B_1) + A_1\overline{p}_1 + B_1(1-\overline{p}_1) = s\ .$$

The expression on $\overline{p}_1$ can be simplified to $A_1(1+\alpha)\overline{p}_1 + B_1\alpha(1-\overline{p}_1) = s(\overline{p}_1+\alpha)$, which combined with $A_1 = \frac{\lambda_1(rh+s)}{r+\lambda_1}$, $B_1 = \frac{\lambda_2 s}{r+\lambda_2}$, and $\alpha = \frac{r+\lambda_2}{\lambda_1-\lambda_2}$ given by Lemma 1, produces the unique solution $\overline{p}_1 = \frac{s}{\lambda_1 h} = p_1^m$. On the experimentation region, u_1 is strictly convex, since $u_1''(p) > 0$ is implied by $C_1 > 0$, as shown below. The value function is strictly decreasing below the cutoff $\overline{p}_1$ and strictly increasing above the cutoff $\overline{p}_1$. Thus, value-matching and smooth-pasting conditions admit a unique solution for $\overline{p}_1$.

The constant of integration can be recovered from the smooth-pasting condition. Given that $\overline{p}_1 = p_1^m$, the odds ratio at the cutoff belief is $\Omega(\overline{p}_1) = \frac{\lambda_1 h - s}{s}$ and

$$C_1 \left(\frac{\lambda_1 h - s}{s}\right)^{\alpha} = \frac{\overline{p}_1}{\overline{p}_1 + \alpha}(A_1 - B_1) \, .$$

Straightforward algebra yields

$$\frac{\overline{p}_1}{\overline{p}_1 + \alpha} = \frac{(\lambda_1 - \lambda_2)s}{(\lambda_1 - \lambda_2)s + \lambda_1(r + \lambda_2)h} \, , \qquad A_1 - B_1 = \frac{r((\lambda_1 - \lambda_2)s + \lambda_1(r + \lambda_2)h)}{(r + \lambda_1)(r + \lambda_2)} \, .$$

Thus, we get the desired equation recovering C_1:

$$C_1 \left(\frac{\lambda_1 h - s}{s}\right)^{\alpha} = \frac{r(\lambda_1 - \lambda_2)s}{(r + \lambda_1)(r + \lambda_2)} \, .$$

Finally, $C_1 > 0$ as $\lambda_1 h > s$ and $\lambda_1 - \lambda_2 > 0$. ■

Proof of ***Lemma 2***:

Proof. The consistency of the agent 2's value function with agent 1's strategy, $u_2(\overline{p}_1) = s$, together with $\overline{p}_1 = p_1^m$ by Proposition 4 require

$$C_2(1 - p_1^m)(\Omega(p_1^m))^{\alpha} = s - A_2 p_1^m - B_2(1 - p_1^m) \, .$$

First, the right-side can be simplified to

$$s - A_2 p_1^m - B_2(1 - p_1^m) = r \cdot \frac{(r + \lambda_1)h(\lambda_1 + \lambda_2 s - \lambda_1 \lambda_2 h) - (\lambda_1 - \lambda_2)s^2}{(r + \lambda_1)(r + \lambda_2)\lambda_1 h} \, .$$

Next, characterization of agent 1's constant of integration in Proposition 4 yields:

$$(1 - p_1^m)(\Omega(p_1^m))^{\alpha} = \frac{\lambda_1 h - s}{\lambda_1 h} \cdot \frac{r(\lambda_1 - \lambda_2)s}{(r + \lambda_1)(r + \lambda_2)} \cdot \frac{1}{C_1} \, .$$

Combining these observations, we have

$$C_2 = -\frac{C_1}{\lambda_1 h - s}\left[\frac{(r + \lambda_1)h(\lambda_1 \lambda_2 h - (\lambda_1 + \lambda_2)s)}{(\lambda_1 - \lambda_2)s} + s\right] \, .$$

Note that since $(\lambda_1 + \lambda_2)s < (\lambda_1 + \lambda_2)\frac{\lambda_2 h}{2} = \frac{\lambda_1 \lambda_2 h}{2} + \frac{\lambda_2^2 h}{2} < \lambda_1 \lambda_2 h$, we have $\lambda_1 \lambda_2 h - (\lambda_1 + \lambda_2)s > 0$. Finally, since we also have $\lambda_1 h - s > 0$ and $C_1 > 0$ by Proposition 4, we conclude that $C_2 < 0$. ■

Lemma 3 *As $\Delta_\lambda = \lambda_1 - \lambda_2$ converges to zero from the right, both constants of integration converge to zero:* $\lim_{\Delta_\lambda \to 0+} C_1 = 0$ *and* $\lim_{\Delta_\lambda \to 0+} C_2 = 0$*.*

Proof of ***Lemma 3****:*

Proof. Fix λ_2 and let $\lambda_1 = \lambda_2 + \Delta_\lambda$, with $\Delta_\lambda \to 0+$.

From Proposition 4, we have

$$C_1 = \frac{rs(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)} \left(\frac{s}{\lambda_1 h - s}\right)^{\alpha} ,$$

where $\alpha = \frac{r+\lambda_2}{\lambda_1 - \lambda_2}$. Since $\lambda_1 h > 2s$, we have $\frac{s}{\lambda_1 h - s} < 1$. Further, $\alpha = \frac{r+\lambda_2}{\Delta_\lambda} \to +\infty$ as $\Delta_\lambda \to 0+$. Thus, we can write $C_1 = \Delta_\lambda G(\Delta_\lambda)$, where $G(\Delta_\lambda) \to 0$ as $\Delta_\lambda \to 0+$. The limit of C_1 then follows.

From Lemma 2,

$$C_2 = -\frac{C_1}{\lambda_1 h - s} \left[\frac{h(r+\lambda_1)\big(\lambda_1\lambda_2 h - s(\lambda_1+\lambda_2)\big)}{s(\lambda_1 - \lambda_2)} + s\right] .$$

Thus, we have

$$|C_2| \leq \frac{|C_1|}{\lambda_1 h - s}\left(\frac{k_1}{\Delta_\lambda} + k_2\right) = \frac{1}{\lambda_1 h - s}\left(k_1|G(\Delta_\lambda)| + k_2 \Delta_\lambda |G(\Delta_\lambda)|\right)$$

for some finite positive constants k_1 and k_2. Since the right-hand side of the inequality above converges to zero as $\Delta_\lambda \to 0+$, the limit of C_2 follows. ■

Proof of ***Proposition 5****:*

Proof. The proof is divided into separate proofs for parts (i) - (iii):

Part (i). Note that u_2 is continuous and strictly concave by Lemma 2. Let P_2 be the set of beliefs at which agent 2 is willing to start the agreement, that is, the set of beliefs p that satisfy $u_2(p) \geq s$. Since u_2 is continuous and strictly concave, P_2 is an interval. Since $u_2(p_1^m) = s$, P_2 is nonempty. Denote $\overline{p}_2$ as a supremum of P_2, so that $u_2(\overline{p}_2) = s$. We conclude that agent 2 is willing to start the agreement if and only if $p_0 \leq \overline{p}_2$, where $\overline{p}_2$ is the maximal solution of $u_2(p) = s$.

Part (ii). Consider agent 2's value when the agreement is in place at $t = 0$ and runs until the stopping time τ either through a breakthrough or through agent 1's exit at p_1^m following no news. This value for starting belief p_0 can be written as follows:

$$u_2(p_0) = s + \mathbb{E}\left[\int_0^\tau re^{-rt}(\lambda_2(1-p_t)h - s)dt\right] . \tag{3}$$

Suppose $p_0 = p_2^m$ and note that $\lambda_2(1-p_0)h = s$. Given the downward drift of p_t under no breakthroughs when agents are asymmetric, we have $\lambda_2(1-p_t)h > \lambda_2(1-p_0)h$ along no-news path. Thus, the integrand in equation (3) is strictly positive for $t \in [0, \tau]$. Since no-news path of positive length realizes with strictly positive probability, we conclude that $u_2(p_2^m) > s$ when $\Delta_\lambda > 0$. Given continuity and strict concavity of u_2 and part (i), we can then conclude that $p_2^m < \overline{p}_2$.

Fix λ_2 and let $\lambda_1 = \lambda_2 + \Delta_\lambda$. Define $f(p, \Delta_\lambda) = u_2(p; \lambda_1 = \lambda_2 + \Delta_\lambda, \lambda_2) - s$. We want to study the cutoff $\overline{p}_2(\Delta_\lambda)$ that satisfies $f(\overline{p}_2(\Delta_\lambda), \Delta_\lambda) = 0$. From the analysis of symmetric agents we know that $\overline{p}_2(\Delta_\lambda = 0) = p_2^m$, $f(p_2^m, 0) = 0$, and $f_p(p_2^m, 0) < 0$. From Lemma 3, $u_2(p; \lambda_1 = \lambda_2 + \Delta_\lambda, \lambda_2)$ is continuously differentiable near $(p_2^m, 0)$. Then by the implicit function theorem, for small Δ_λ there is a unique continuous function $\tilde{p}_2(\Delta_\lambda)$ that solves $f(p, \Delta_\lambda) = 0$ in a neighborhood of p_2^m. We already showed that for each $\Delta_\lambda > 0$ equation $f(p, \Delta_\lambda) = 0$ has a unique solution greater than p_2^m. Therefore, we must have $\overline{p}_2(\Delta_\lambda) = \tilde{p}_2(\Delta_\lambda)$. By continuity, we thus have $\overline{p}_2(\Delta_\lambda) \to p_2^m$ as $\Delta_\lambda \to 0+$.

Part (iii). First, note that since $\lambda_2 h > 2s$, we have $p_2^m > \frac{1}{2}$ and hence $\bar{p}_2 > \frac{1}{2}$ by part (ii). $\bar{p}_2$ satisfies $u_2(\bar{p}_2; \lambda_1) = s$. By the implicit function theorem,

$$\frac{\partial \bar{p}_2}{\partial \lambda_1} = -\frac{\partial u_2(\bar{p}_2; \lambda_1)/\partial \lambda_1}{\partial u_2(\bar{p}_2; \lambda_1)/\partial p} \, .$$

Note that $\frac{\partial u_2(\bar{p}_2;\lambda_1)}{\partial p} < 0$ by concavity of u_2 and part (i). To analyze the sign of $\frac{\partial u_2(\bar{p}_2;\lambda_1)}{\partial \lambda_1}$, we can define the following functions that are parts of agent 2's value function:

$$\begin{aligned} \mathcal{A}(p, \lambda_1) &\equiv C_2(1-p)(\Omega(p))^\alpha \, , \\ \mathcal{B}(p, \lambda_1) &\equiv A_2 p + B_2(1-p) \, , \end{aligned}$$

Then

$$\frac{\partial u_2(\bar{p}_2; \lambda_1)}{\partial \lambda_1} = \frac{\partial \mathcal{A}(\bar{p}_2; \lambda_1)}{\partial \lambda_1} + \frac{\partial \mathcal{B}(\bar{p}_2; \lambda_1)}{\partial \lambda_1} \, ,$$

where we can directly calculate:

$$\begin{aligned} \frac{\partial \mathcal{A}}{\partial \lambda_1} &= (1-p)(\Omega(p))^\alpha \left[\frac{\partial C_2}{\partial \lambda_1} - \frac{C_2(r+\lambda_2)\log(\Omega(p))}{(\lambda_1-\lambda_2)^2} \right] \, , \\ \frac{\partial \mathcal{B}}{\partial \lambda_1} &= \frac{psr}{(r+\lambda_1)^2} > 0 \, . \end{aligned}$$

Since $(1/\Omega(p))^\alpha$, the exponential function of α, grows faster than the polynomial function of α, we have $\frac{\partial \mathcal{A}}{\partial \lambda_1} \to 0$ as $\Delta_\lambda \to 0+$ (and $\alpha \to \infty$). Indeed, applying the implicit function theorem to $u_2(\bar{p}_1; \lambda_1) = s$, we can show that

$$\begin{aligned} \frac{\partial C_2}{\partial \lambda_1} = &- C_2 \left[\frac{s}{\lambda_1(\lambda_1 h - s)} + \frac{r+\lambda_2}{(\lambda_1-\lambda_2)\Omega(\bar{p}_1)} \frac{h}{s} - \frac{r+\lambda_2}{(\lambda_1-\lambda_2)^2} \log(\Omega(\bar{p}_1)) \right] + \\ &\left(\frac{\lambda_2(rh+s)s}{(r+\lambda_2)h\lambda_1^2} - \frac{s^2}{(r+\lambda_1)^2 h} \right) \frac{1}{(1-\bar{p}_1)(\Omega(\bar{p}_1))^\alpha} \, . \end{aligned}$$

Then as $\Delta_\lambda \to 0+$, we have

$$\frac{\partial \mathcal{A}}{\partial \lambda_1} \to -\, C_2(1-p)(\Omega(p))^\alpha \left[\frac{s}{\lambda_1(\lambda_1 h - s)} + \frac{r+\lambda_2}{(\lambda_1-\lambda_2)\Omega(\bar{p}_1)}\frac{h}{s} - \frac{r+\lambda_2}{(\lambda_1-\lambda_2)^2}\log(\Omega(\bar{p}_1)) + \frac{(r+\lambda_2)\log(\Omega(p))}{(\lambda_1-\lambda_2)^2}\right] \sim \nu C_2(\Omega(p))^\alpha \alpha^2 \ ,$$

where ν is a scalar. Applying L'Hopital's rule twice, $(\Omega(p))^\alpha\alpha^2 \to 0$ as $\alpha \to \infty$ whenever $p > 1/2$. This implies

$$\lim_{\Delta_\lambda \to 0+} \frac{\partial \bar{p}_2}{\partial \lambda_1} \ > \ 0 \ .$$

By continuity (part (ii)), it must be that $\frac{\partial \bar{p}_2}{\partial \lambda_1} > 0$ when λ_1 is sufficiently close to λ_2.

Following similar steps, we can show that $\frac{\partial \bar{p}_2}{\partial \lambda_2} > 0$ when λ_1 and λ_2 are sufficiently close. Indeed, by the implicit function theorem,

$$\frac{\partial \bar{p}_2}{\partial \lambda_2} = -\frac{\partial u_2(\bar{p}_2;\lambda_2)/\partial \lambda_2}{\partial u_2(\bar{p}_2;\lambda_2)/\partial p} \ .$$

As before, note that $\frac{\partial u_2(\bar{p}_2;\lambda_2)}{\partial p} < 0$ by concavity of u_2 and part (i). To analyze the sign of $\frac{\partial u_2(\bar{p}_2;\lambda_2)}{\partial \lambda_2}$, we can define the following functions that are parts of agent 2's value function with a slight abuse of notation:

$$\begin{aligned} \mathcal{A}(p,\lambda_2) &\equiv C_2(1-p)(\Omega(p))^\alpha \ , \\ \mathcal{B}(p,\lambda_2) &\equiv A_2 p + B_2(1-p) \ , \end{aligned}$$

Then

$$\frac{\partial u_2(\bar{p}_2;\lambda_2)}{\partial \lambda_2} \ = \ \frac{\partial \mathcal{A}(\bar{p}_2;\lambda_2)}{\partial \lambda_2} + \frac{\partial \mathcal{B}(\bar{p}_2;\lambda_2)}{\partial \lambda_2} \ ,$$

where we can directly calculate:

$$\begin{aligned} \frac{\partial \mathcal{A}}{\partial \lambda_2} &= (1-p)(\Omega(p))^\alpha \left[\frac{\partial C_2}{\partial \lambda_2} + \frac{C_2(r+\lambda_1)\log(\Omega(p))}{(\lambda_1-\lambda_2)^2}\right] \ , \\ \frac{\partial \mathcal{B}}{\partial \lambda_2} &= \frac{(1-p)(rh+s)r}{(r+\lambda_2)^2} > 0 \ . \end{aligned}$$

Thus, $\frac{\partial \mathcal{A}}{\partial \lambda_2} \to 0$ $\Delta_\lambda \to 0+$ (and $\alpha \to \infty$). Indeed, applying the implicit function theorem to $u_2(\bar{p}_1;\lambda_2) = s$, we can show that

$$\frac{\partial C_2}{\partial \lambda_2} = -C_2\frac{r+\lambda_1}{(\lambda_1-\lambda_2)^2}\log(\Omega(\bar{p}_1)) - \frac{r(rh+s)}{(r+\lambda_2)^2}\frac{1}{(\Omega(\bar{p}_1))^\alpha} \ .$$

Then as $\Delta_\lambda \to 0+$, we have

$$\frac{\partial \mathcal{A}}{\partial \lambda_2} \to -C_2(1-p)(\Omega(p))^\alpha \frac{r+\lambda_1}{(\lambda_1-\lambda_2)^2}\left[\log(\Omega(\bar{p}_1)) - \log(\Omega(p))\right] \sim \nu C_2 (\Omega(p))^\alpha \alpha^2,$$

where ν is a scalar. Applying L'Hopital's rule twice, $(\Omega(p))^\alpha \alpha^2 \to 0$ as $\alpha \to \infty$ whenever $p > 1/2$. This implies

$$\lim_{\Delta_\lambda \to 0+} \frac{\partial \bar{p}_2}{\partial \lambda_2} \quad > \quad 0 \; .$$

By continuity, it must be that $\frac{\partial \bar{p}_2}{\partial \lambda_2} > 0$ when λ_2 is sufficiently close to λ_1. ■

Lemma 4 *When agents are asymmetric, the sum of agents' constants of integration is negative, that is,* $C_1 + C_2 < 0$*.*

Proof of ***Lemma 4****:*
Proof. By Lemma 2, we have

$$C_1 + C_2 = C_1\left(1 - \frac{1}{\lambda_1 h - s}\left[\frac{(r+\lambda_1)h[\lambda_1\lambda_2 h - s(\lambda_1+\lambda_2)]}{(\lambda_1-\lambda_2)s} + s\right]\right) \; ,$$

where $C_1 > 0$ by Proposition 4. Thus, we want to show that

$$\frac{(r+\lambda_1)h[\lambda_1\lambda_2 h - s(\lambda_1+\lambda_2)]}{(\lambda_1-\lambda_2)s} > \lambda_1 h - 2s \; .$$

Note that if this inequality holds for $r = 0$, then it holds for all $r > 0$. Thus, we want to show that

$$\lambda_1 h[\lambda_1\lambda_2 h - s(\lambda_1+\lambda_2)] > (\lambda_1 h - 2s)(\lambda_1 - \lambda_2)s \; .$$

Expanding the brackets, we get

$$\begin{aligned}
\lambda_1^2\lambda_2 h^2 - \lambda_1^2 sh - \lambda_1\lambda_2 sh > \lambda_1^2 sh - 2\lambda_1 s^2 - \lambda_1\lambda_2 sh + 2\lambda_2 s^2 &\iff \\
\lambda_1^2\lambda_2 h^2 > 2\lambda_1^2 sh - 2\lambda_1 s^2 + 2\lambda_2 s^2 &\iff \\
\lambda_1\lambda_2 h(\lambda_1 h - s + s) > 2\lambda_1 s(\lambda_1 h - s) + 2\lambda_2 s^2 &\iff \\
\lambda_1\lambda_2 h(\lambda_1 h - s) + \lambda_1\lambda_2 hs > 2\lambda_1 s(\lambda_1 h - s) + 2\lambda_2 s^2 & \; ,
\end{aligned}$$

where the last inequality holds because $\lambda_1 h > \lambda_2 h > 2s$. Thus, we obtain the desired result. ■
Proof of ***Proposition 6****:*

Proof. Fix asymmetric λ_1, λ_2, with $\Delta_\lambda > 0$. Then for the agreement value, we have

$$V'(p) = \frac{r(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)}(rh + 2s) - (C_1 + C_2)\frac{\alpha + p}{p}(\Omega(p))^\alpha > 0 \, ,$$
$$V''(p) = (C_1 + C_2)\frac{1}{p^2}\left(\alpha + \frac{\alpha + p}{1-p}\right)(\Omega(p))^\alpha < 0 \, .$$

The signs of the derivatives follow directly from Lemma 4. Thus, V is increasing and strictly concave on the experimentation region.

For the societal value, we have

$$S'(p) = \frac{r(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)}(-b) + C_S\frac{\alpha + p}{p}(\Omega(p))^\alpha \, ,$$

where

$$C_S = \frac{1}{(1-p_1^m)(\Omega(p_1^m))^\alpha}\left(\frac{1-p_1^m}{r+\lambda_2} + \frac{p_1^m}{r+\lambda_1}\right) rb > 0 \, .$$

Then S is strictly concave on the experimentation region since

$$S''(p) = -C_S\frac{1}{p^2}\left(\alpha + \frac{\alpha + p}{1-p}\right)(\Omega(p))^\alpha < 0 \, .$$

Furthermore, we can show that $S'(p_1^m) > 0$. Indeed, $S'(p_1^m) = \frac{rb}{\Delta_\lambda p_1^m(1-p_1^m)} > 0$ follows directly from evaluating ODE guiding the societal value at p_1^m and using boundary condition $S(p_1^m) = 0$. Note that this implies $W'(p_1^m) > 0$.

We can conclude that $W(p) = (1-\beta)V(p) + \beta S(p)$ is strictly concave on the experimentation region, and thus, there is a unique welfare-maximizing prior p^* lying in the experimentation region.

If $S'(\overline{p}_2) \geq 0$, then $W(p)$ is increasing on the experimentation region and thus is maximized at $p = \overline{p}_2$. We now show that if $\frac{\lambda_i h}{s}$ is sufficiently large, then it has to be the case that $S'(\overline{p}_2) < 0$.

We note that

$$\left(\frac{1-p_1^m}{r+\lambda_2} + \frac{p_1^m}{r+\lambda_1}\right) = \frac{r + \lambda_1 - \frac{s}{\lambda_1 h}(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)} \, .$$

Then

$$\left(\frac{1-p_1^m}{r+\lambda_2} + \frac{p_1^m}{r+\lambda_1}\right) \cdot \frac{1}{1-p_1^m} = \frac{(r+\lambda_1)\lambda_1 h - s(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)(\lambda_1 h - s)} \, ,$$

and

$$S'(p) = rb\left[-\frac{(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)} + \frac{(r+\lambda_1)\lambda_1 h - s(\lambda_1 - \lambda_2)}{(r+\lambda_1)(r+\lambda_2)(\lambda_1 h - s)}\left(\frac{\alpha}{p} + 1\right)\left(\frac{\Omega(p)}{\Omega(p_1^m)}\right)^\alpha\right] \, .$$

We want to show that at $p = \overline{p}_2$:

$$\frac{(r+\lambda_1)\lambda_1 h - s(\lambda_1 - \lambda_2)}{(\lambda_1 h - s)}\left(\frac{\alpha}{p} + 1\right)\left(\frac{\Omega(p)}{\Omega(p_1^m)}\right)^\alpha < \lambda_1 - \lambda_2 \, .$$

Using the fact that $\alpha = \frac{r+\lambda_2}{\lambda_1-\lambda_2}$ and $\alpha+1 = \frac{r+\lambda_1}{\lambda_1-\lambda_2}$, we get

$$\frac{(\alpha+1)\lambda_1 h - s}{(\lambda_1 h - s)}\left(\frac{\alpha}{p}+1\right)\left(\frac{\Omega(p)}{\Omega(p_1^m)}\right)^{\alpha} < 1 \; .$$

We now note that the left-hand side is decreasing in p. Thus, it suffices to show that this inequality is satisfied at $p = p_2^m < \overline{p}_2$ by Proposition 5. At $p = p_2^m$, we have $\frac{\alpha}{p_2^m}+1 = \frac{(\alpha+1)\lambda_2 h - s}{\lambda_2 h - s}$. Furthermore, $\frac{\alpha}{1-p_1^m}+1 = \frac{(\alpha+1)\lambda_1 h - s}{\lambda_1 h - s}$. Thus,

$$\left(\frac{\alpha}{1-p_1^m}+1\right)\left(\frac{\alpha}{p_2^m}+1\right)\left(\frac{\Omega(p_2^m)}{\Omega(p_1^m)}\right)^{\alpha} < 1 \; .$$

Taking logarithms of both sides, we get

$$\log\left(\frac{\alpha}{1-p_1^m}+1\right) + \log\left(\frac{\alpha}{p_2^m}+1\right) + \alpha\log\left(\frac{\Omega(p_2^m)}{\Omega(p_1^m)}\right) < 0 \; .$$

Define function $g(\alpha)$ defined on the nonnegative real numbers as the left-hand side of the inequality above. Then we can see that $g(0) = 0$, $\lim_{\alpha\to\infty} g(\alpha) = -\infty$, and $g''(\alpha) < 0$. Then $g'(0) < 0$ produces the desired result. Note that

$$g'(0) = \frac{1}{1-p_1^m} + \frac{1}{p_2^m} + \log\left(\frac{\Omega(p_2^m)}{\Omega(p_1^m)}\right) \; .$$

Substituting the values of myopic cutoffs, we get

$$g'(0) = \frac{\lambda_1 h}{\lambda_1 h - s} + \frac{\lambda_2 h}{\lambda_2 h - s} + \log\left(\frac{s^2}{(\lambda_1 h - s)(\lambda_2 h - s)}\right) \; ,$$

or

$$g'(0) = \frac{s}{\lambda_1 h - s} + \frac{s}{\lambda_2 h - s} + \log\left(\frac{s^2}{(\lambda_1 h - s)(\lambda_2 h - s)}\right) + 2 \; .$$

Thus, if $\frac{s}{\lambda_i h - s} + \log\left(\frac{s}{\lambda_i h - s}\right) < -1$, we obtain the desired result. This inequality is true for $\frac{s}{\lambda_i h - s} < 0.278$, or $\frac{\lambda_i h}{s} > 4.6$. Thus, we obtain a sufficient condition, an explicit lower bound on $\frac{\lambda_i h}{s}$, under which $S'(\overline{p}_2) < 0$. Suppose this sufficient condition holds. Then $W(p) = (1-\beta)V(p) + \beta S(p)$ cannot be maximized at p_1^m as $W'(p_1^m) > 0$. Furthermore, $W(p)$ is maximized at $\overline{p}_2$ if and only if $(1-\beta)V'(\overline{p}_2) + \beta S'(\overline{p}_2) \geq 0$, or $\beta \leq \frac{V'(\overline{p}_2)}{V'(\overline{p}_2) - S'(\overline{p}_2)} := \hat{\beta} \in (0,1)$. The solution is thus interior if $\beta > \hat{\beta}$. This completes the proof. ■

Proof of ***Proposition*** *7:*

Proof. If $p_0 \notin [p_1^m, \overline{p}_2]$, then there is nothing to prove as the agreement does not start in the asymmetric case. Consider $p_0 \in [p_1^m, \overline{p}_2]$. For fixed parameters r, s, h, β, b, define the following function:

$$\omega(\lambda) = (1-\beta)\frac{\lambda}{r+\lambda}(rh+2s) + \beta\frac{1}{r+\lambda}rb \; .$$

Then the affine portion of the asymmetric welfare $W(p_0; \lambda_1, \lambda_2)$ is $p_0\omega(\lambda_1) + (1-p_0)\omega(\lambda_2)$. Furthermore, $W(p_0; \lambda_1, \lambda_2) < p_0\omega(\lambda_1) + (1-p_0)\omega(\lambda_2)$ when $\lambda_1 > \lambda_2$ by Lemma 4 and a negative asymmetric drift term in the societal value. By continuity of $\omega(\lambda)$ and the intermediate-value theorem, for $p_0 \in [0, 1]$, there always exists $\hat{\lambda} \in [\lambda_2, \lambda_1]$, such that $\omega(\hat{\lambda}) = p_0\omega(\lambda_1) + (1-p_0)\omega(\lambda_2)$. But $\omega(\hat{\lambda}) = W(p_0, \hat{\lambda}, \hat{\lambda})$ if p_0 lies in the experimentation region for symmetric case $\hat{\lambda}$. Thus, if p_0 lies in the experimentation region for symmetric case $\hat{\lambda}$, then we provided a welfare-improving symmetrization of λ_1 and λ_2. If p_0 does not lie in the experimentation region for symmetric case $\hat{\lambda}$, then further moving $\hat{\lambda}$ to either $\overline{\lambda}(p_0) = \max\left\{\frac{s}{p_0 h}, \frac{s}{(1-p_0)h}\right\}$ or $+\infty$ can only improve the symmetric welfare and also ensure that the agreement starts at p_0 (Proposition 2 and Proposition 3). This completes the proof. ■

Appendix B: Derivation of societal value $S(p)$ for asymmetric case

The original differential equation is given by

$$\Delta_\lambda p(1-p)S'(p) + (\lambda_1 p + \lambda_2(1-p) + r)S(p) = rb \; .$$

The corresponding homogeneous equation is

$$\Delta_\lambda p(1-p)S'(p) + (\lambda_1 p + \lambda_2(1-p) + r)S(p) = 0 \; .$$

We use separation of variables to solve the homogeneous equation:

$$\begin{aligned}
\Delta_\lambda \frac{dS}{dp} &= -\left(\frac{\lambda_1 + r}{1-p} + \frac{\lambda_2 + r}{p}\right) S \; , \\
\Delta_\lambda \log(S) &= (\lambda_1 + r)\log(1-p) - (\lambda_2 + r)\log(p) + \tilde{\tilde{C}}, \tilde{\tilde{C}} \in \mathbb{R} \; , \\
\log(S) &= \log(1-p)^{\frac{\lambda_1+r}{\Delta_\lambda}} + \log(p)^{-\frac{\lambda_2+r}{\Delta_\lambda}} + \tilde{C}, \tilde{C} \in \mathbb{R} \; .
\end{aligned}$$

We then use variation of parameters to solve the original differential equation:

$$\begin{aligned}
S(p) &= f(p)(1-p)^{\frac{\lambda_1+r}{\Delta_\lambda}} p^{-\frac{\lambda_2+r}{\Delta_\lambda}} \; , \\
S'(p) &= f'(p)(1-p)^{\frac{\lambda_1+r}{\Delta_\lambda}} p^{-\frac{\lambda_2+r}{\Delta_\lambda}} - \\
& f(p)\left[\frac{\lambda_1 + r}{\Delta_\lambda}(1-p)^{\frac{\lambda_2+r}{\Delta_\lambda}} p^{-\frac{\lambda_2+r}{\Delta_\lambda}} + \frac{\lambda_2 + r}{\Delta_\lambda}(1-p)^{\frac{\lambda_1+r}{\Delta_\lambda}} p^{-\frac{\lambda_1+r}{\Delta_\lambda}}\right] \; .
\end{aligned}$$

Substituting these into the original nonhomogeneous differential equation gives

$$\Delta_\lambda f'(p) p^{\frac{\lambda_1-2\lambda_2-r}{\Delta_\lambda}}(1-p)^{\frac{2\lambda_1-\lambda_2+r}{\Delta_\lambda}} = rb \; , \tag{4}$$

$$\frac{\Delta_\lambda}{rb} f'(p) = p^{\frac{\lambda_2+r}{\Delta_\lambda}-1}(1-p)^{-\frac{\lambda_1+r}{\Delta_\lambda}-1} \; . \tag{5}$$

Next, we calculate

$$\int p^{\frac{\lambda_2+r}{\Delta_\lambda}-1}(1-p)^{-\frac{\lambda_1+r}{\Delta_\lambda}-1} dp = \Delta_\lambda (1-p)^{-\frac{\lambda_1+r}{\Delta_\lambda}} p^{\frac{\lambda_2+r}{\Delta_\lambda}} \left(\frac{1}{r+\lambda_2}(1-p) + \frac{1}{r+\lambda_1}p\right) + C, C \in \mathbb{R}$$

Combining with equation (5), we get

$$f(p) = (1-p)^{-\frac{\lambda_1+r}{\Delta_\lambda}} p^{\frac{\lambda_2+r}{\Delta_\lambda}} \left(\frac{r}{r+\lambda_2}(1-p) + \frac{r}{r+\lambda_1}p\right) b + \hat{C}, \hat{C} \in \mathbb{R} \; .$$

Since, $S(p) = f(p)(1-p)^{\frac{\lambda_1+r}{\Delta_\lambda}} p^{-\frac{\lambda_2+r}{\Delta_\lambda}}$, we get

$$S(p) = \left(\frac{1}{r+\lambda_2}(1-p) + \frac{1}{r+\lambda_1}p\right) rb + \hat{C}(1-p)^{\frac{\lambda_1+r}{\Delta_\lambda}} p^{-\frac{\lambda_2+r}{\Delta_\lambda}} \ .$$

Now, using the boundary condition $S(p_1^m) = 0$, we can pin down $\hat{C}$:

$$\hat{C} = -\frac{\frac{1-p_1^m}{r+\lambda_2} + \frac{p_1^m}{r+\lambda_1}}{(1-p_1^m)^{\frac{\lambda_1+r}{\Delta_\lambda}} (p_1^m)^{-\frac{\lambda_2+r}{\Delta_\lambda}}} \cdot rb \ ,$$

so that finally

$$S(p) = \left(\frac{1-p}{r+\lambda_2} + \frac{p}{r+\lambda_1}\right) rb - \left(\frac{1-p_1^m}{r+\lambda_2} + \frac{p_1^m}{r+\lambda_1}\right) rb \left(\frac{1-p}{1-p_1^m}\right)^{\frac{\lambda_1+r}{\Delta_\lambda}} \left(\frac{p}{p_1^m}\right)^{-\frac{\lambda_2+r}{\Delta_\lambda}} \ .$$

Appendix C: Strictly increasing societal value $S(p)$ for low agreement stakes

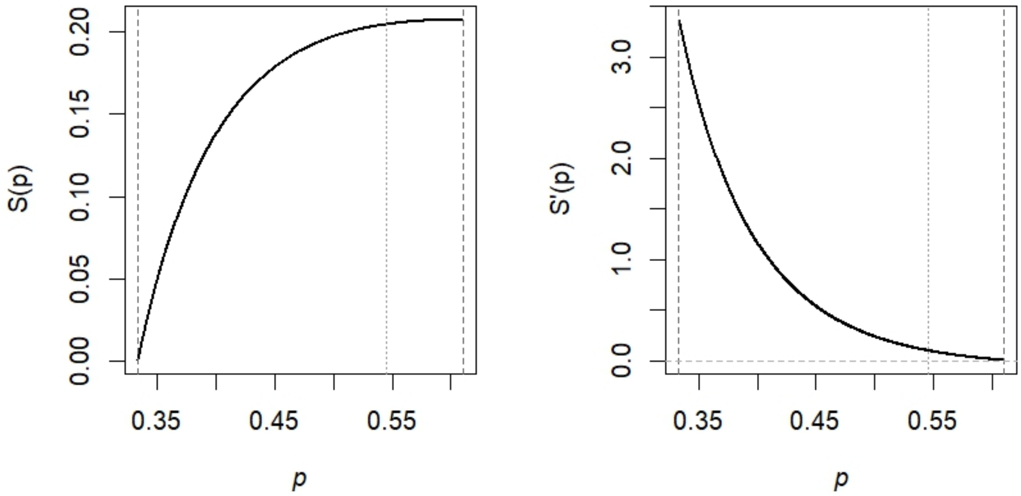

Figure 5: Societal value $S(p)$ (left panel) and marginal societal value $S'(p)$ (right panel) on the experimentation region for $\lambda_1 = 0.75$, $\lambda_2 = 0.55$, $b = 3$, $s = 0.5$, $h = 2$, $r = 0.05$. The vertical dashed lines from left to right correspond to p_1^m, p_2^m, and $\overline{p}_2$.

Figure 5 illustrates that in the environment of low agreement stakes the societal value can be strictly increasing in the experimentation region. In particular, under the specific parameters of Figure 5, $S'(\overline{p}_2) \approx 0.004$. Therefore, the welfare-maximizing prior is $p^* = \overline{p}_2$.