Dynamics of Risky Agreements

Renee Bowen Malte Lammert Aleksandr Levkun

May 22, 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. To model these features we assume each agent has access to a risky arm of a Poisson bandit and a safe outside option. The agreement is in force only if *both* agents are pulling their respective risky arms. Risky arms are negatively correlated, reflecting the fact that only one agent is favored by the agreement. If one agent quits, then the agreement is dissolved and both agents receive their safe payoff. Agents will enter the agreement only if they are sufficiently optimistic that they are favored and will quit when sufficiently convinced they are not favored. Agreements are, thus, in force for an interval of beliefs. This implies some efficient agreements are never started, and efficient agreements in force end with certainty. We find that under some conditions welfare is non-monotonic in the agreement's informativeness, and is maximized for interior priors. This suggests that slow judgments and agreement ambiguity can improve welfare under some conditions. We further show that a mild asymmetry between the speed of judgments for each player, can improve welfare for some starting beliefs.

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 judgement. 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 about 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. 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). 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 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 judgement and we consider agent 1 to be the agent with the higher frequency of judgments condition on having a good arm. Exactly one agent has the good arm and one agent has the bad arm, and the agent with the good arm is ex-ante unknown. The bandit arms are thus negatively correlated as in

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. The belief that agent 2 has the good risky arm is simply 1 minus agent 1's belief because of the perfect negative correlation. 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 good 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 halts for all agents. We study equilibria in cutoff 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 these 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. Agents simply weigh the safe outside payoff against the expected return from the agreement, given their current belief. 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 it is straightforward to show that agreement value increases with the frequency of judgments, as agents prefer to receive the uncertain reward sooner. In contrast, 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.

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. In the asymmetric case, the *level* of the frequency of judgments matters for welfare as in the symmetric case, but, in addition the *difference* between the two agents' judgment frequency matters. We show that a mild asymmetry between agents may be beneficial for welfare for a range of initial beliefs.

The remainder of this paper is structured as follows: Section 2 reviews related literature, Section 3 introduces the formal model, Section 4 characterizes the equilibrium in cut-off strategies as well as the myopic benchmark beliefs, Section 5 and 6 lay out the analysis for symmetric and asymmetric agents, respectively and Section 7 concludes.

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), which is mainly concerned with exponential bandits. A bandit is a probabilistic device that generates observable evidence in accordance to pre-determined 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 freeriding 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 reveals the state. Keller and Rady (2010) study positive correlation when news are inconclusive. If bandits are negatively correlated, the freeriding incentive is reversed, so that experimentation is encouraged. Klein and Rady (2011) study negatively correlated bandits in environment where news are conclusive. They organize their findings within three categories, depending on the stakes, i.e. potential gains from the risky arm. 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 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, agents act myopically. Most relevant to our analysis, Keller and Rady (2015) study a model of positive correlation and inconclusive news with the novelty that news are more likely to occur in an undesired state, so that upon observation of news, beliefs jump downwards rather than upwards and eventually experimentation breaks down. We study a model of negative correlation and conclusive news with the novelty that if one agent stops experimenting, experimentation stops for all agents. To our knowledge, such interdependence has not been studied yet, but 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.

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 agreementual 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) show that a seller can benefit from an ambiguous buying mechanism wich hides certain features of the mechanism when the buyer is ambiguity averse. Grimmelmann (2019) shows that such ambiguity can even be found in smart agreements written in programming languages. 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 and infinite, i.e. $t \in [0, \infty)$. 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 payoffs are uncertain when in force. 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 $j \neq i$. 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 higher than if the agreement favors agent 2. 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.

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[1]

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

[1] The probability of a reward arriving for both agents or multiple rewards arriving for the same agent is negligible as we are ignoring terms of order $o(dt)$.

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 to experiment in order 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_{-i}. 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.[2] We henceforth refer to a Markov equilibrium simply as an equilibrium.

4 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 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 $\overline{p}_1 \leq p \leq \overline{p}_2$. By standard arguments, while the agreement is in force agents' value

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

functions satisfy the 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 following ODE:

$$dp = -p(1-p)\Delta\lambda dt$$

where $\Delta\lambda = \lambda_1 - \lambda_2$ captures the asymmetry in the agreement. Generally, as $\Delta\lambda > 0$, 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. Thus learning has a positive impact on agent 1's agreement value, but a negative impact on agent 2's. 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, to the expected instantaneous payoff to being in the agreement—$\lambda_1 ph$, for player 1 and $\lambda_2(1-p)h$ for player 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 $\lambda_2 h > 2s$ implies $p_1^m < p_2^m$, 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 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. For a prior belief close to middle of the experimentation range, there is a longer period of learning, so the learning speed can influence the longevity of the agreement. We begin to explore how these two forces interact with a simple case of symmetric agents.

5 Symmetric Agents

To develop some intuition we begin by assuming the distribution of rewards conditional on p is identical for both agents, i.e., $\lambda_1 = \lambda_2 = \lambda$. The first implication is that with equal arrival rates, if no reward arrives, beliefs remain constant, i.e. $dp = 0$. The reason is that while for agent i no payoff arriving is bad news, no payoff arriving for agent $-i$ is good news of the exact same magnitude. Notice that informativeness $\Delta\lambda$ and the speed of learning in the absence of rewards is zero. Simplifying agents' value functions (1) and (2) we obtain the linear equations

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

These have straightforward solutions

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

Note $u_1(p)$ 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 own pessimism. Agent 1's threshold

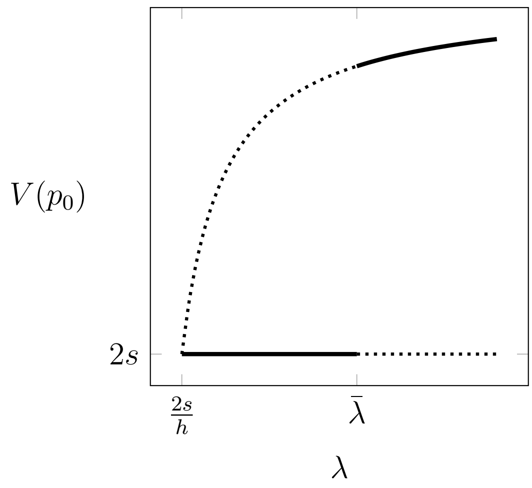

Figure 1: Agreement value $V(p_0)$ for fixed p_0 as a function of λ.

belief thus satisfies $u_1(\overline{p}_1) = s$, while agent 2's implies $u_2(\overline{p}_2) = s$. These conditions yield the agents' myopic beliefs as their threshold beliefs. We summarize this 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* $\overline{p}_1 \le p \le \overline{p}_2$*, where* $\overline{p}_1 = p_1^m$ *and* $\overline{p}_2 = p_2^m$ *.*

The equilibrium in Proposition 1 describes an agreement that starts if and only if myopic agents would want to enter the agreement, that is $p_1^m < p_0 < p_2^m$ and ends when there is a breakthrough for either player. This happens with probability λdt over time interval dt. Notice that since p_1^m is decreasing in λ while p_2^m is increasing in λ, a higher λ allows entry into the agreement for a larger set of priors.

Agreement value We define the *agreement value* V as the sum of agents' expected payoffs given an initial belief p_0. If the agents initiate the agreement at the initial belief p_0, that is, $p_1^m \le p_0 \le p_2^m$, 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 started, and the agreement value is equal to a sum of safe payoffs. Therefore,

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

Notice that if $p_1^m \le p_0 \le p_2^m$ then $\lim_{\lambda \to \frac{2s}{h}} V(p_0) = 2s$ and $\lim_{\lambda \to \infty} V(p_0) = rh + 2s$. Define $\bar{\lambda} = \max\left\{\frac{s}{p_0 h}, \frac{s}{(1-p_0)h}\right\}$ as the cutoff on λ, such that if $\lambda < \bar{\lambda}$, fixing other parameters, one of the players prefer to exit. Figure 1 shows the agreement value for fixed prior belief p_0 across different λ.

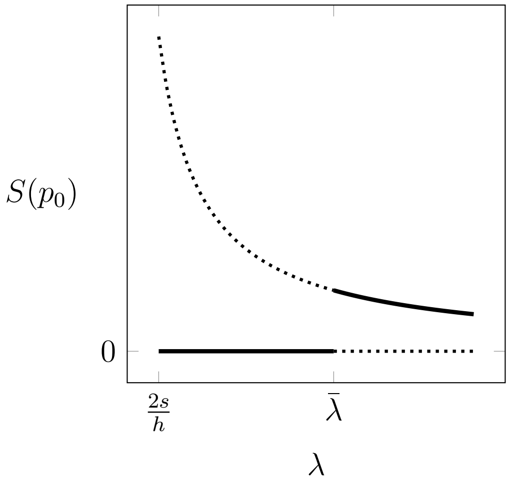

Figure 2: Societal value $V(p_0)$ for fixed p_0 as a function of λ.

Societal value Assume that the agreement, while in force, exerts a flow externality b on society. We define the *societal value* S as the expected present discounted value of the stream of flow externalities given an initial belief p_0. Given that the agreement is in force, the societal value is

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

If the agreement is not in force, the societal value is zero, as no externalities arrive. Therefore,

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

For values of λ below $\bar{\lambda}$, the agreement ends immediately without realizing any societal value. As long as the agreement starts, the societal value decreases in λ as the agreement stays in place shorter, with $\lim_{\lambda\to\infty} S(p_0) = 0$. Figure 2 shows the societal value for fixed prior belief p_0 across different λ.

Welfare Welfare is a linear combination of the agreement value and the societal value:

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

As a result, we have:

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

If $p_1^m \le p_0 \le p_2^m$, we can rewrite:

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

Then for the maximization problem $W(p_0) \to \max_\lambda$, the solution is

$$\lambda^* = \begin{cases} +\infty & \text{if } (1-\beta)(rh+2s) - \beta b \ge 0 \\ \bar{\lambda} & \text{if } (1-\beta)(rh+2s) - \beta b \le 0 \end{cases}$$

Note that the myopic cutoffs are functions of λ as well, so the above solution has the following implications for welfare beginning with the case when societal value has a low weight in total welfare.

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

1. *Instant judgments are optimal.*
2. *When judgments are instant the value of 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 high societal value is quite different, and depends on agents' prior belief.

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

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

Figure 3 illustrates the optimal value as a function of the prior belief for the case of a relatively high weight placed on societal value. With slow optimal judgments, the welfare achieves maximum across prior beliefs when the prior belief does not favor any agent, $p_0 = 1/2$.

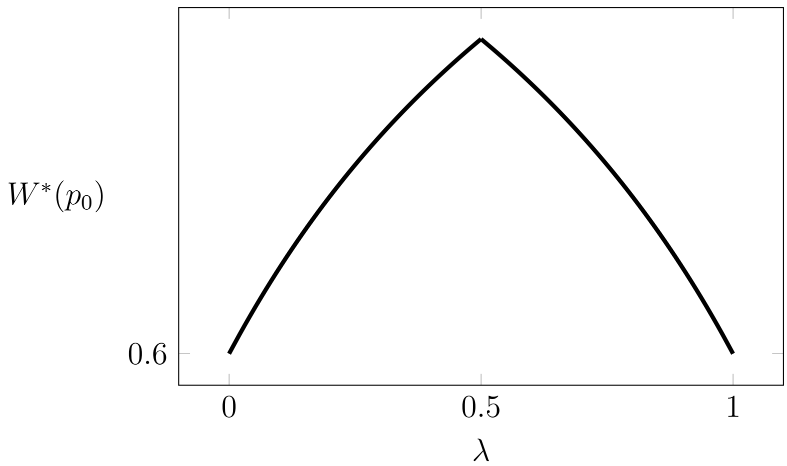

Figure 3: The value function for welfare $W^*(p_0)$ exemplified for $\beta = 0.5, b = 3, s = 0.5, h = 2, r = 0.1$.

6 Asymmetric Agents

To understand the role of belief drift, we first characterize the equilibrium cutoff beliefs when agents are asymmetric, i.e. $\lambda_1 > \lambda_2$. Notice that since $\Delta\lambda > 0$, the agreement is informative and beliefs drift down while the agreement is in force and no rewards arrive for any agent.

Lemma 1 *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 $\Omega(p) = \frac{1-p}{p}$ *and* A_i, B_i *and* C_i *are constants.*

$\Omega(p) = \frac{1-p}{p}$ is the odds ratio at belief p. The term $C_1(1-p)(\Omega(p))^\alpha$ is the option value of being able to exit the agreement, while the term $A_1 p + B_1(1-p)$ is the expected payoff of committing to the agreement. 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.

Substituting the guess into agent 1's value function in (1) and matching coefficients we obtain

$$A_1 = \frac{\lambda_1(rh+s)}{r+\lambda_1}, \quad B_1 = \frac{\lambda_2 s}{r+\lambda_2}, \quad \alpha = \frac{r+\lambda_2}{\lambda_1-\lambda_2}.$$

Intuitively, we have $A_1 > s > B_1$. Note also that the linear part of the value function is the same as in the case of $\lambda_1 = \lambda_2$, since the expected payoff of committing to the agreement is not affected by the drift.[3]

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

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

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

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

To see why agent 1's threshold is the same as the myopic threshold, note that the arrival of a reward for agent 1 immediately kills 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 value of experimentation as a result. Only myopic incentives are relevant as a result. This will not be the case for agent 2. Observe that agent 1's constant of integration C_1 is positive and value function u is convex and achieves a minimum at $\overline{p}_1$. Agent 1's value is thus decreasing over time while the agreement is in force. Also note that

$$\lim_{\Delta\lambda \to 0} C_1 = 0,$$

confirming our intuition above that the option value of experimenting goes to zero with symmetric agents.

We now turn to agent 2. Due to the drift of the belief process, for a given prior p_0 agent 2 either never enters the agreement or participates until the first reward arrives. As with agent 1, we guess that agent 2's value function takes the form

$$u_2(p) = C_2(1 - p)(\Omega(p))^{\alpha} + A_2 p + B_2(1 - p) \ .$$

We substitute this guess into agent 2's value function in (2) an match linear coefficients to obtain

$$A_2 = \frac{\lambda_1 s}{r + \lambda_1}, \quad B_2 = \frac{\lambda_2(rh + s)}{r + \lambda_2}. \tag{3}$$

As before, the linear portion of agent 2's value function collapses to the value function for agent 2 in the symmetric case.

What remains is to determine the constant of integration C_2 and the threshold belief. 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, it must also be true that when agent 1 decides to quit the agreement, agent 2's value falls to s. That is, it must satisfy continuity of u_2 given the strategy of agent 1 —$u_2(\overline{p}_1) = s$. This latter condition pins down agent 2's constant of integration.

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

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

Lemma 2 and the equations in (3) complete 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. Consistent with the previous intuition,

$$\lim_{\Delta\lambda\to 0} C_2 = 0,$$

We illustrate both agents' equilibrium value functions below along with the equilibrium cutoffs.

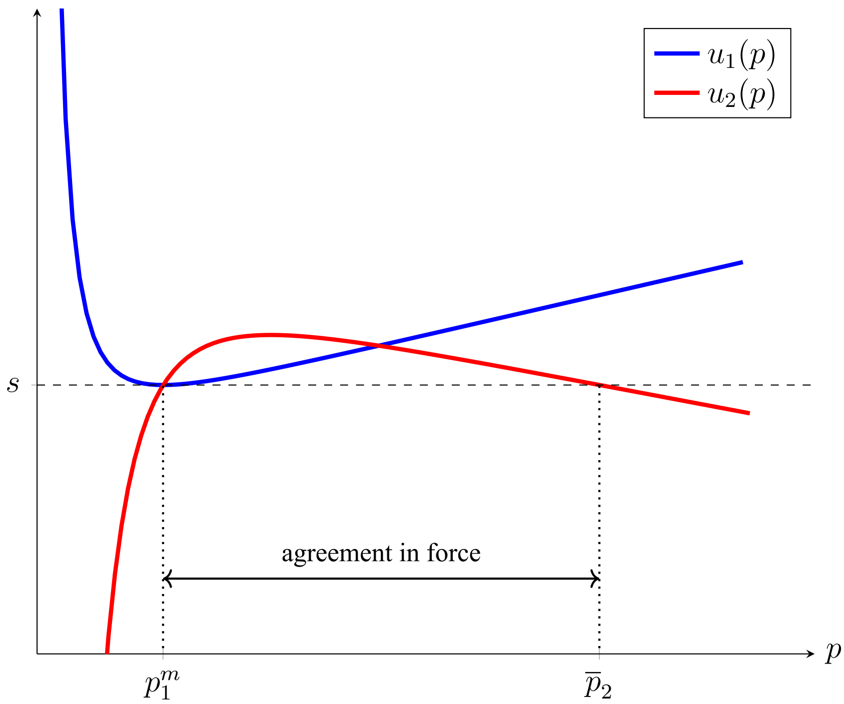

Figure 4: Agents' value functions and equilibrium cutoffs.

Intuitively, $u_2(p)$ is equal to s at $\overline{p}_2$, and the value of C_2 is such that $u_2(p)$ also attains the value s on the left at $\overline{p}_1$. We formally characterize $\overline{p}_2$ in the following proposition.

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

(i) agent 2 chooses to start the agreement for $p_0 \le \overline{p}_2$, 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) for $\Delta\lambda$ sufficiently small, $\overline{p}_2 > p_2^m$.

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

Proof. See Appendix. ■

In the symmetric case studied above the agreement was dissolved only by a breakthrough of either agent 1 or 2. In the asymmetric case, now that there is a drift of beliefs, in addition the agreement may be dissolved if for a sufficiently long time no rewards arrive. Hence, fixing the total probability of a breakthrough, i.e. $p\lambda_1 + (1-p)\lambda_2$, an asymmetric agreement is generally shorter lived than a symmetric one. Therefore, opportunity cost, i.e. foregoing a safe payoff of s for the duration of the agreement being in force is lower for both agents. This motivates player 2 to stay in the agreement beyond p_2^m. In addition, increasing λ_1 increases $\Delta\lambda$ and thus the speed of learning, decreasing overall longevity even further. Therefore, the higher λ_1, the more agent 2 is willing to experiment beyond her myopic cutoff p_2^m.

Agreement value As in the case of symmetric agents, the agreement value is given by the sum of agents' value functions, $u_1(p_0)+u_2(p_0)$, if the agreement is in place for belief p_0. Otherwise, the agreement collapses, and the agreement value is equal to a sum of 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

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

where the constants of integration, C_1 and C_2, are given by Proposition 4 and Lemma 2, respectively.

Societal value 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 $S(p_1^m) = 0$. Solving this boundary-value problem, we obtain[4]

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

when $p_1^m \leq p \leq \overline{p}_2$; and $S(p) = 0$, otherwise.

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

Welfare Welfare is a linear combination of the agreement value and the societal value:

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

To outline the effect of asymmetry on welfare, define an increasing function $\rho(\mu)$ for fixed parameter λ, such that if $\lambda_1 = \lambda + \mu$ and $\lambda_2 = \lambda - \rho(\mu)$, then $\bar{p}_2 = p_2^m$. That is, an increase in λ_1 by μ and a corresponding decrease of λ_2 by $\rho(\mu)$ preserves agent 2's incentives to initiate the agreement. Function ρ is well-defined as agent 2's cutoff is monotonic in λ_1 and λ_2.

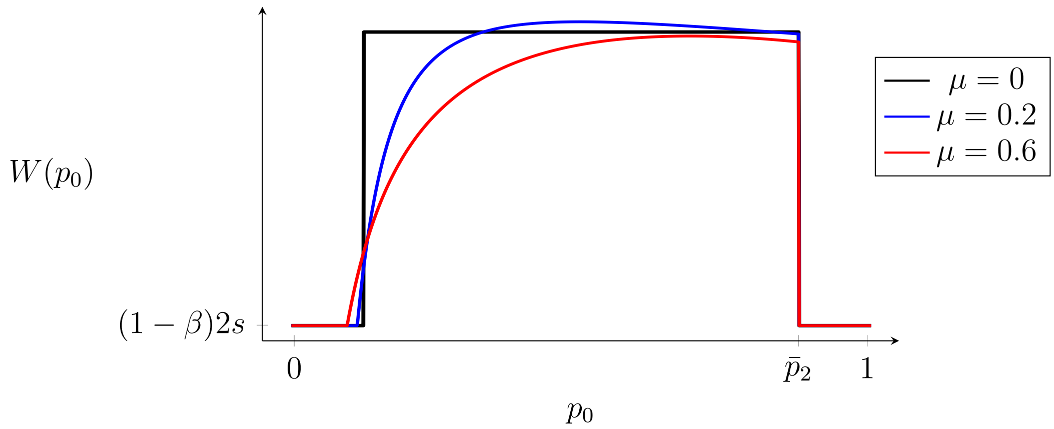

Figure 5: Welfare function for $\lambda_1 = \mu$, $\lambda_2 = \lambda - \rho(\mu)$, $\beta = 0.5, b = 3, s = 0.5, h = 2, r = 0.1$.

Figure 5 illustrates that an introduction of mild asymmetry between agents may be beneficial for welfare for a range of initial beliefs.

7 Concluding Remarks

We use strategic experimentation to model the dynamics of risky agreements. The novelty of our approach is that payouts require coordination, that is, if one agent stops to experiment, i.e. quits the agreement, experimentation stops for all. 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 one agent. Longevity of an agreement depends on the evolution of beliefs and therefore on the probability of news arriving as well as on the informativeness of an agreement in the absence of news (in the asymmetric case). An agreement dissolves if beliefs evolve too extreme in any one direction and trigger one agent to quit. We show that agreements are generally inefficient. Many agreements, although efficient from 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. Once initiated, an agreement's longevity is governed by the intensity of experimentation as well as the informativeness. Generally, more informative agreements with faster learning can lead to longer or shorter agreements, depending on the stake of the agreement. The same holds true for welfare. More informative agreements can increase or decrease welfare, depending on stakes.

8 Appendix A: Proofs

Proof of ***Proposition 4***:

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

Part (i). The proof follows from the main text.

Part (ii). First, as $v(1/2;\lambda_1) > s$ whenever λ_1 and λ_2 are close to each other, we have $\bar{p}_2 > 1/2$ by concavity of v. $\bar{p}_2$ satisfies $v(\bar{p}_2;\lambda_1) = s$. By the implicit function theorem

$$\frac{\partial \bar{p}_2}{\partial \lambda_1} = -\frac{\partial v/\partial \lambda_1}{\partial v/\partial p}.$$

Define

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

Then

$$\begin{aligned}
v(\bar{p}_2;\lambda_1) &= \mathcal{A}(\bar{p}_2,\lambda_1) + \mathcal{B}(\bar{p}_2,\lambda_1) \\
\frac{\partial v}{\partial \lambda_1} &= \frac{\partial \mathcal{A}}{\partial \lambda_1} + \frac{\partial \mathcal{B}}{\partial \lambda_1} \\
\frac{\partial v}{\partial p} &= \frac{\partial \mathcal{A}}{\partial p} + \frac{\partial \mathcal{B}}{\partial p}
\end{aligned}$$

We can calculate

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

Since $(1/\Omega(p))^{\alpha}$, the exponential function of α, grows faster than the polynomial function of α, we have $\frac{\partial \mathcal{A}}{\partial \lambda_1} \to 0$ and $\frac{\partial \mathcal{A}}{\partial p} \to 0$ as $\lambda_1 \to \lambda_2$ (or $\alpha \to \infty$). Indeed, applying the implicit function theorem to $v(\bar{p}_1;\lambda_1)=s$, we can show that

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

Then as $\lambda_1 \to \lambda_2$, 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 $p > 1/2$. Similarly, by applying L'Hopital's rule once, $\frac{\partial \mathcal{A}}{\partial p} \to 0$ as $\alpha \to \infty$. This implies

$$\lim_{\lambda_1 \to \lambda_2} \frac{\partial \bar{p}_2}{\partial \lambda_1} > 0.$$

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

Part (iii). The proof follows directly from part (ii). To see that, notice that when $\lambda_1 = \lambda_2$, $\bar{p}_2 = p_2^m$. Hence, when $\bar{p}_2$ increases in λ_1, it must be that $\bar{p}_2 > p_2^m$ for $\lambda_1 > \lambda_2$. ■

WTS: $\frac{\partial \bar{p}_2}{\partial \lambda_2} > 0$ when λ_1 and λ_2 are sufficiently close.
By the implicit function theorem

$$\frac{\partial \bar{p}_2}{\partial \lambda_2} = -\frac{\partial v/\partial \lambda_2}{\partial v/\partial p}.$$

Define

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

Then

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

We can calculate

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

Since $(1/\Omega(p))^\alpha$, the exponential function of α, grows faster than the polynomial function of α, we have $\frac{\partial \mathcal{A}}{\partial \lambda_1} \to 0$ and $\frac{\partial \mathcal{A}}{\partial p} \to 0$ as $\lambda_1 \to \lambda_2$ (or $\alpha \to \infty$). Indeed, applying the implicit function theorem to $v(\bar{p}_1; \lambda_2) = s$, 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 $\lambda_2 \to \lambda_1$, 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 $p > 1/2$. Similarly, by applying L'Hopital's rule once, $\frac{\partial \mathcal{A}}{\partial p} \to 0$ as $\alpha \to \infty$. This implies

$$\lim_{\lambda_2 \to \lambda_1} \frac{\partial \bar{p}_2}{\partial \lambda_2} > 0.$$

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

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

The original differential equation is given by

$$\Delta\lambda p(1-p)S'(p) + (\lambda_1 p + \lambda_2(1-p) + r)L(p) = rb$$

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:

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

Substituting these into the original nonhomogeneous differential equation gives

$$\begin{aligned}
\Delta\lambda f'(p) p^{\frac{\lambda_1 - 2\lambda_2 - r}{\Delta\lambda}}(1-p)^{\frac{2\lambda_1 - \lambda_2 + r}{\Delta\lambda}} &= rb \\
\frac{\Delta\lambda}{rb} f'(p) &= p^{\frac{\lambda_2+r}{\Delta\lambda}-1}(1-p)^{-\frac{\lambda_1+r}{\Delta\lambda}-1} \qquad (*)
\end{aligned}$$

Next, we calculate

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

Using this in (*), 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

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

10 Appendix C: Derivation of $L(p)$ for asymmetric case

The original differential equation is given by

$$\Delta\lambda p(1-p)L'(p) + (\lambda_1 p + \lambda_2(1-p))L(p) = 1$$

The corresponding homogeneous equation is

$$\Delta\lambda p(1-p)L'(p) + (\lambda_1 p + \lambda_2(1-p))L(p) = 0$$

We use separation of variables to solve the homogeneous equation:

$$\begin{aligned}
\Delta\lambda \frac{\partial L}{\partial p} &= -\left(\frac{\lambda_1}{1-p} + \frac{\lambda_2}{p}\right) L \\
\Delta\lambda \log(L) &= \lambda_1 \log(1-p) - \lambda_2 \log(p) + \tilde{\tilde{C}}, \tilde{\tilde{C}} \in \mathbb{R} \\
\log(L) &= \log(1-p)^{\frac{\lambda_1}{\Delta\lambda}} - \log(p)^{\frac{-\lambda_2}{\Delta\lambda}} + \tilde{C}, \tilde{C} \in \mathbb{R}
\end{aligned}$$

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

$$\begin{aligned}
L(p) &= f(p)(1-p)^{\frac{\lambda_1}{\Delta\lambda}} p^{\frac{-\lambda_2}{\Delta\lambda}} \\
L'(p) &= f'(p)(1-p)^{\frac{\lambda_1}{\Delta\lambda}} p^{\frac{-\lambda_2}{\Delta\lambda}} - f(p)\left[\frac{\lambda_1}{\Delta\lambda}(1-p)^{\frac{\lambda_2}{\Delta\lambda}} p^{\frac{-\lambda_2}{\Delta\lambda}} + \frac{\lambda_2}{\Delta\lambda}(1-p)^{\frac{\lambda_1}{\Delta\lambda}} p^{\frac{-\lambda_1}{\Delta\lambda}}\right]
\end{aligned}$$

Substituting these into the original differential equation gives

$$\begin{aligned}
\Delta\lambda f'(p) p^{\frac{\lambda_1 - 2\lambda_2}{\Delta\lambda}} (1-p)^{\frac{2\lambda_1 - \lambda_2}{\Delta\lambda}} &= 1 \\
\Delta\lambda f'(p) &= p^{\frac{\lambda_2}{\Delta\lambda} - 1}(1-p)^{\frac{-\lambda_2}{\Delta\lambda} - 2} \qquad (*)
\end{aligned}$$

Next, we calculate

$$\int p^{\frac{\lambda_2}{\Delta\lambda} - 1}(1-p)^{\frac{-\lambda_2}{\Delta\lambda} - 2} dp = \frac{\Delta\lambda}{\lambda_1\lambda_2}(1-p)^{\frac{-\lambda_1}{\Delta\lambda}} p^{\frac{\lambda_2}{\Delta\lambda}}(\lambda_1(1-p) + \lambda_2 p) + C, C \in \mathbb{R}$$

Using this in (*), we get

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

Since, $L(p) = f(p)(1-p)^{\frac{\lambda_1}{\Delta\lambda}} p^{\frac{-\lambda_2}{\Delta\lambda}}$, we get

$$L(p) = \frac{1-p}{\lambda_2} + \frac{p}{\lambda_1} + \hat{C}(1-p)^{\frac{\lambda_1}{\Delta\lambda}} p^{\frac{-\lambda_2}{\Delta\lambda}}$$

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

$$\hat{C} = -\frac{\frac{1-p_1^m}{\lambda_2} + \frac{p_1^m}{\lambda_1}}{(1-p_1^m)^{\frac{\lambda_1}{\Delta\lambda}} {p_1^m}^{\frac{-\lambda_2}{\Delta\lambda}}}$$

so that finally

$$L(p) = \frac{1-p}{\lambda_2} + \frac{p}{\lambda_1} - \left(\frac{1-p_1^m}{\lambda_2} + \frac{p_1^m}{\lambda_1}\right)\left(\frac{1-p}{1-p_1^m}\right)^{\frac{\lambda_1}{\Delta\lambda}}\left(\frac{p}{p_1^m}\right)^{\frac{-\lambda_2}{\Delta\lambda}}$$

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.

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).

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.

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.

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).