Greener Thy Neighbor? On the Welfare Effects of Protectionist Climate Policies

Saumya Deojain David Lindequist

December 16, 2024

Abstract

The world is witnessing a surge in green industrial policies, with prominent examples such as the U.S. Inflation Reduction Act (IRA) incorporating significant protectionist elements. While traditionally viewed as distortive, we argue that in the case of climate policies, protectionism can have strategic value by facilitating coordination between countries on climate action. We present a simple model that blends a standard abatement game with a zero-sum beggar-thy-neighbor game, allowing for the possibility of multiple equilibria. Using techniques from the global games literature, we show that uncertainty surrounding the distortions caused by protectionist policies yields a unique equilibrium. We find that protectionist climate policies improve welfare when expected distortions are low, as they promote coordination on climate change mitigation at relatively modest costs. High expected distortions make these policies welfare-neutral, as countries are unlikely to adopt them. For intermediate expected distortions, protectionist policies are most harmful, combining a high probability of coordination failure with substantial costs. Our findings suggest that regulators like the WTO could boost global welfare by restricting but not entirely banning protectionist climate policies. We further find that efforts to reduce uncertainty about distortions might backfire, inadvertently encouraging coordination on costly protectionist measures.

JEL Classification: C72, D83, Q56

Keywords: climate policy coordination, protectionism, beggar-thy-neighbor policies, global games, inflation reduction act, coordination games

---

We thank Christoph Carnehl, Josh Ederington, Doug Hanley, Ahmad Lashkaripour, Georg Schaur, and various seminar audiences for useful comments. All errors are our own.

Deojain: Thapar School of Liberal Arts and Sciences, and University of Rochester, W. Allen Wallis Institute of Political Economy; e-mail: saumya.deojain@thapar.edu.

Lindequist: Miami University; e-mail: lindeqd@miamioh.edu

1. Introduction

Industrial policy has reemerged as a central topic in economic debates, with countries actively promoting industries critical to national security and the green transition toward sustainability (Juhász, Lane, and Rodrik 2023; Mazzucato and Rodrik 2023). In the United States, the CHIPS and Science Act of 2022 aims to strengthen semiconductor research and manufacturing, reflecting this renewed focus. This revival has been extensively analyzed, particularly regarding global semiconductor markets (Goldberg et al. 2024; Bown and Wang 2024) and broader protectionist trends (Fajgelbaum et al. 2020).

The Inflation Reduction Act (IRA) of 2022, considered the most ambitious climate legislation in U.S. history (Bistline et al. 2023), directs $783 billion toward domestic energy security and climate initiatives. Positioned as a leading example of green industrial policy (Altenburg and Rodrik 2017), approximately $369 billion of this funding is earmarked for green subsidies, including tax credits and direct financial support. Notably, many of these subsidies apply exclusively to products made in North America through 'domestic content requirements.' For instance, $12 billion is allocated for electric vehicle incentives, but only for vehicles assembled in North America. Additionally, the IRA imposes tariffs on certain green technologies, such as solar cells and lithium-ion batteries, to protect and promote the U.S. green technology sector. This strategy, coupled with incentives for domestic production, has been characterized as a shift toward green protectionism (Attinasi, Boeckelmann, and Baptiste 2023), exemplifying the intersection of environmental objectives with industrial policy goals (Rodrik 2014).

These protectionist measures have drawn criticism from several U.S. trading partners. The European Union (EU), in particular, has been vocal, even considering legal action against the U.S. at the World Trade Organization (WTO) or contemplating retaliatory protectionist measures (Fajeau et al. 2023). This has raised concerns about a potential 'green arms race,' where countries engage in zero-sum competition to attract international firms and capital (Kleimann et al. 2023; Gros, Mengel, and Presidente 2023; Dahlström et al. 2023; Sattich and Huang 2023).

Economists have largely expressed discomfort with the protectionist components of the IRA. Historically, they have been wary of industrial policies due to risks like the negative welfare consequences of beggar-thy-neighbor protectionism (Evenett 2019), the potential for rent-seeking, and the challenges in selecting winning industries (Rodrik 2020). There are also concerns about trade frictions resulting from such policies (Irwin 2011), which could distort international supply chains and labor markets, ultimately leading to welfare losses (Canayaz, Erel, and Gurun 2024).

In this paper, we theoretically examine the welfare effects of protectionist climate policies, highlighting a crucial distinction from standard protectionist measures. While beggar-thy-neighbor (BTN) policies—zero-sum strategies where one country's gain directly results in another's loss—typically create distortions leading to inefficiencies, these distortions may, in the context of the green transition, mitigate the lack of coordination in climate policies and enhance social welfare. Specifically, BTN elements in climate policy increase the desirability of mitigation in a climate game that would otherwise suffer from the underprovision of a global public good.

We argue that provisions encouraging the domestic production of a public good—such as those in the IRA—can help overcome this global underprovision by promoting coordination on the green transition, even though they distort the efficient allocation of international capital. This contrasts with policies like the CHIPS Act, where support for the domestic production of non-public goods is essentially zero-sum and offers no increase in global welfare. Although the WTO currently restricts industrial policies that distort trade, particularly subsidies, we suggest that in a second-best world, these restrictions may be detrimental to global welfare. Trade distortions could provide an effective pathway to tackling climate change.

Our work builds on the long-standing insight from the theory of the second best, originally outlined by Lipsey and Lancaster (1956). This theory posits that if eliminating a specific market distortion—such as externalities in public good provision—is impractical, introducing an additional distortion in related markets, like BTN frictions in international trade, may offset the original distortion, potentially resulting in a more efficient outcome. The central question driving this paper is under what conditions protectionist climate policy can actually improve global welfare.

To clarify the relevant trade-offs and structure our analysis of the welfare effects of protectionist green policy, we model two identical countries deciding whether to undergo the green transition—a binary action choice. Each country's decision generates a positive externality for the other, as both benefit from reduced global emissions. Social welfare is maximized when both countries transition, but since the costs are privately borne and countries do not fully internalize the global benefits, they individually find it optimal not to transition. In this 'pure climate game,' not transitioning is the dominant strategy, resulting in a prisoner's dilemma where neither country transitions.

We extend this model to allow countries to implement protectionist policies during the green transition, introducing what we call the 'BTN climate game.' Here, countries can allocate a portion of their transition effort to BTN policies, which we term BTN effort. In this game, a country can extract a resource transfer from the other if its BTN effort exceeds that of the other. However, BTN effort incurs a cost, denoted by κ, reflecting welfare losses from supply chain distortions and the rising cost of goods caused by protectionist measures.

First, we consider the case where BTN costs κ are known. If these costs are sufficiently low, both countries will find it optimal to transition with full protectionism. For high values of κ, the dominant strategy is not to transition. At intermediate values of κ, the BTN climate game exhibits multiple equilibria: a country will transition if the other does and will not transition if the other does not. This multiplicity stems from the convexity of payoffs in the number of countries transitioning. At intermediate BTN costs, transitioning is attractive only if the other country transitions, creating a coordination problem and resulting in multiple equilibria.

A fundamental issue in the real world is that the welfare costs from BTN-induced trade distortions are uncertain. We introduce this uncertainty by assuming both countries share a common prior belief about the true BTN cost parameter before making their transition decision. Each country then receives a private signal about the cost and updates its belief accordingly. This framework leads to a global game in the sense of Carlsson and Van Damme (1993). We demonstrate that under mild conditions, the game with uncertain BTN costs has a unique equilibrium in switching strategies.[1] If countries believe BTN costs are low, they will transition; if they believe the costs are high, they will not. As is typical in global games, this uncertainty combined with sufficiently precise private signals eliminates the multiplicity of equilibria found in the coordination game. The unique equilibrium provides a probability distribution over the possible payoffs of the BTN climate game, allowing us to derive comparative statics for expected welfare based on the model parameters.

We find that the expected welfare in the unregulated BTN climate game is non-monotonic in the mean of the prior belief about the BTN cost parameter κ. When expected BTN costs are very low, both countries will almost certainly transition with full protectionism, leading to welfare that exceeds the payoff from the pure climate game. In this case, it is socially efficient to endure BTN costs for the sake of the green transition. When BTN costs are very high, neither country will transition, which is also efficient, as the costs of trade distortions outweigh the benefits of transitioning. However, for intermediate BTN costs, welfare can be lower than in the pure climate game. This occurs for one of two reasons. First, if welfare improves when both countries transition but not when only one does, intermediate BTN costs create a high probability that only one country will transition,

[1]These conditions reflect those in Morris and Shin (2000) and require the signal about the BTN costs to be sufficiently precise relative to the precision of the prior belief about these costs.

reducing welfare relative to the pure climate game. Second, if expected BTN costs are just high enough to make full protectionist transitions undesirable, the likelihood of both countries transitioning remains relatively high, resulting in a welfare loss compared to the pure climate game.

While the BTN climate game can lead to a welfare improvement over the pure climate game, the zero-sum nature of BTN transfers means that more protectionist policy is implemented in equilibrium than is necessary to incentivize transitions. We show that a regulator who controls the level of protectionist policy can improve expected welfare. By reducing the extent of BTN policies in the BTN game, a regulator can increase the welfare benefits over the pure climate game, provided expected BTN costs are not excessively high.

In terms of policy conclusions, our results suggest that the WTO should be restrictive toward protectionist climate policies when their expected distortions are of intermediate size, as this creates a high risk of costly coordination failures, making the use of such policies too expensive for facilitating the green transition. However, when expected distortions are low, the WTO should permit some protectionist measures to support the green transition. In contrast, when expected distortions are high, strict WTO regulation becomes less critical, as countries are unlikely to adopt protectionist climate policies due to their prohibitive costs.

In practice, no global regulatory authority has the power to control protectionist policies fully. However, institutions like the WTO (or the economics profession more broadly) can offer more precise information about the expected distortions these policies cause.[2] We find that the welfare effects of such information are complex. While greater precision can increase the likelihood of countries undergoing a green transition, it may also prompt this transition in cases where protectionist distortions are high, leading to suboptimal outcomes. In these instances, less precise information may actually improve welfare by preventing undesirable coordination on costly protectionist measures.

In summary, our paper takes a first step toward understanding the global welfare effects of green protectionist policies by analyzing their strategic value in facilitating global climate action. Protectionist climate policies, like those in the IRA, may enhance international cooperation on climate action. However, costly coordination failures remain a real threat, especially when the expected costs of trade distortions are moderate, complicating the job of international regulators like the WTO.

[2] Fajgelbaum and Khandelwal (2022), Attinasi, Boeckelmann, and Meunier (2023), Balistreri and McDaniel (2024), and Robinson and Thierfelder (2024) are examples of the booming literature trying to assess the welfare costs of the recent reemergence of distortive trade policies.

1.1. Related Literature

Our work contributes to the vast literature on international climate change policy coordination. The overview articles by Finus (2000), Barrett (1994, 2005), and Wood (2011) highlight the challenges of achieving stable cooperation in international environmental agreements due to strong free-rider incentives. In related work, Fischer (2017) explores how interventionist green policies, often seen as distorting, might actually enhance global welfare—particularly in situations where implementing carbon prices is politically challenging, as noted by Juhász and Lane (2024). We extend this literature by introducing protectionist climate policies, demonstrating how these policies can shift strategic incentives and potentially foster cooperation.

Our work adds to the reemerging literature on trade policy and climate agreements and is closely related to studies by Barrett and Dannenberg (2022) and Andres (2023). Barrett and Dannenberg (2022) show that linking trade agreements to public goods provision enhances cooperation, with multilateral approaches outperforming unilateral ones. Similarly, Andres (2023) finds that temporary protectionism and subsidies can foster clean technology development and yield long-term welfare gains. While these studies also emphasize the effect of industrial policy on climate policy coordination, we incorporate uncertainty about the costs of trade distortions to explore the welfare implications of protectionist climate policies.

Harstad (2024) and Farrokhi and Lashkaripour (2024) focus on the climate effects of institutional mechanisms like contingent trade agreements and carbon border taxes. Unlike their approaches, we highlight how beggar-thy-neighbor protectionist policies can shift incentives for cooperation without formal agreements. Ederington (2010) explores whether trade agreements should include environmental policy, while we suggest that it is the absence of free trade agreements that can help mitigate the underprovision of climate action. Additionally, Hambel, Kraft, and Schwartz (2018) examine carbon abatement dynamics and free-riding, whereas we focus on how trade distortions interact with climate mitigation efforts in a strategic context.

Ulph (2004), Kolstad (2007), and Barrett and Dannenberg (2012) emphasize the role of uncertainty around future costs and benefits in shaping climate agreements. We extend this work by focusing on the uncertainty surrounding economic distortions from protectionist policies, demonstrating how such uncertainty can either lead to welfare-improving coordination or result in costly coordination failures. While applications of global games methodology to climate coordination are rare, notable exceptions include Heijmans (2022) and Heijmans (2023), who explores equilibrium selection in the context of network externalities.[3] Our paper examines how the availability of protectionist policies impacts climate cooperation, which, to the best of our knowledge, has not been studied before.

Zehaie (2009), Bayramoglu, Finus, and Jacques (2018), and Hritonenko, Hritonenko, and Yatsenko (2020) investigate the trade-offs between adaptation and mitigation strategies. In contrast, our work focuses on the interaction between protectionism and mitigation rather than adaptation. By combining protectionist measures with mitigation efforts, we analyze how these policies reshape the strategic environment in international climate cooperation. In studies like Rubio and Ulph (2007), de Zeeuw (2008), Heitzig, Lessmann, and Zou (2011), and Nordhaus (2021), the stability of self-enforcing agreements is explored. We contribute by demonstrating how distortive protectionist measures can introduce strategic incentives that help sustain cooperation. Lastly, Harstad (2012), Nordhaus (2015), and Martimort and Sand-Zantman (2016) focus on the design of optimal climate contracts. While these papers rely on formal agreements and mechanisms, our work shows that protectionist policies can influence incentives for cooperation even without binding contracts, offering a new lens through which to view the effects of protectionism on international climate coordination.

The rest of this paper is organized as follows. Section 2 presents the basic model setup. In Section 3, we analyze the global game resulting from uncertainty about the welfare costs of protectionist policy. We derive our main results regarding the welfare effects of protectionist climate policy in Section 4. Section 5 concludes. All proofs are in Appendix A.

2. Model Setup

Consider two countries, referred to as players. For any player i ('she'), let j ('he') denote the other player, such that $j \neq i$. Each player has a binary abatement ('green transition') choice $a_i = \{0, 1\}$, where $a_i = 0$ signifies 'no abatement' and $a_i = 1$ indicates 'abatement.' Furthermore, each player selects a protectionism level per unit of abatement, denoted by $\alpha_i \in [0, 1]$, where $\alpha_i = 0$ corresponds to no protectionism and $\alpha_i = 1$ represents full protectionism, meaning all abatement expenditure is protectionist. For example, if 30% of climate policy spending by player i is protectionist, then $\alpha_i = 0.3$.

The total level of protectionism by player i is denoted by:

$$\tau_i = \alpha_i a_i = \begin{cases} \alpha_i & \text{if } a_i = 1, \\ 0 & \text{if } a_i = 0. \end{cases} \tag{1}$$

[3] See Morris and Shin (2003), Frankel, Morris, and Pauzner (2003), Jorge and Rocha (2015), and Morris (2016) for an introduction and discussion of global game methodology.

In the event that player i adopts a greater level of protectionism than player j, as indicated by $\tau_i > \tau_j$, player i receives a resource transfer from player j. This setup captures the intent of protectionist policies to favor domestic production at the expense of foreign production or to draw foreign capital. Following Rodrik (2020), we assume a reduced-form zero-sum transfer function of the form:

$$T_i(\tau_i, \tau_j) = \begin{cases} 0 & \text{if } a_i = a_j = 0, \\ \alpha_i & \text{if } a_i = 1, a_j = 0, \\ -\alpha_j & \text{if } a_i = 0, a_j = 1, \\ \alpha_i - \alpha_j & \text{if } a_i = a_j = 1, \end{cases} \tag{2}$$

which implies that $T_i(\tau_i, \tau_j) + T_j(\tau_i, \tau_j) = 0$. Moreover, we assume that 'beggar-thy-neighbor' (BTN) policies incur a private cost that is linearly related to the level of protectionism: $\kappa_i(\tau_i) = \kappa_i \tau_i$ where $\kappa_i \geq 0$. These costs serve as a simplified representation of the distortions stemming from mercantilist policies that lead to efficiency losses in international trade or capital allocation.[4]

Climate mitigation serves as a public good from which both players benefit. Let $\gamma_i u(G)$ represent the utility player i derives from public good $G = a_A + a_B$. The parameter $\gamma_i > 0$ reflects player i's enjoyment of the public good. Alternatively, it represents the damages player i experiences from unmitigated climate change.

We assume that the utility derived from the public good is increasing in its overall provision: $u(2) > u(1) > u(0)$. Crucially, we further assume that the marginal effect of climate change mitigation increases in the number of abating players: $u(2) - u(1) > u(1) - u(0)$. Intuitively, joint efforts lead to more effective results than isolated actions. This assumption reflects the importance of international cooperation in addressing climate change, as the benefits of such cooperation are significantly higher than if each player were to act alone. Relatedly, tipping points in the context of climate change refer to critical thresholds at which a small change in environmental conditions can lead to a significant and often irreversible change in the ecosystem. As the number of players engaged in mitigation increases, the effectiveness of these efforts grows disproportionately. This non-linearity reflects how avoiding tipping points can prevent disproportionately large negative impacts. If fewer players engage in mitigation, the world risks crossing these tipping points, leading to severe and possibly irreversible environmental consequences.

[4]For instance, conditioning subsidies on EVs to require battery production within the US could potentially raise the subsidies required to promote EV purchases compared to a scenario where all EVs receive subsidies, irrespective of battery origin.

Climate policies entail a private cost, denoted by $c_i(a_i)$. For simplicity, assume that the cost function is linear: $c_i(a_i) = c_i a_i$, where $c_i > 0$ indicates player i's private cost of abatement (climate policy implementation). This cost may be high either because a player is relatively poor, making climate change mitigation particularly costly in terms of opportunity costs of resource allocation, or because it is technologically challenging for that player to reduce emissions.

The welfare function of player i is given by:

$$W_i = \gamma_i u(G) - c_i a_i - \kappa_i \tau_i + T_i(\tau_i, \tau_j), \tag{3}$$

where $G = a_i + a_j$ and $\tau_i = \alpha_i a_i$. For the remainder of the analysis, we will assume that players are identical, i.e., $c_i = c_j = c$ and $\gamma_i = \gamma_j = \gamma$. We further make the following crucial assumption:

ASSUMPTION 1. *Let there be $N = 2$ players. For each player it is always privately inefficient but socially efficient to abate. This means that*

$$\gamma \left[u(1) - u(0)\right] < c < 2\gamma \left[u(1) - u(0)\right] \quad \text{and} \quad \gamma \left[u(2) - u(1)\right] < c < 2\gamma \left[u(2) - u(1)\right]. \tag{4}$$

Assuming that $u(2) - u(1) > u(1) - u(0)$, we can succinctly write:

$$\gamma \left[u(2) - u(1)\right] < c < 2\gamma \left[u(1) - u(0)\right]. \tag{5}$$

Furthermore, both players abating is socially efficient. This means that

$$2\gamma u(2) - 2c > 2\gamma u(0), \tag{6}$$

which is guaranteed by assuming that $\gamma \left[u(2) - u(0)\right] > c$.

These assumptions imply that abatement is privately inefficient for each player because the private costs exceed the private benefit, but it is socially efficient as the social benefit outweighs the private cost. This creates a standard collective action problem where countries do not abate despite it being the socially optimal choice. The payoff matrix of the climate game with protectionist policy is presented in Table 1.

2.1. Pure climate game

To fix ideas, first consider a pure climate game without BTN policies. Assumption 1 guarantees that each player's dominant strategy is to choose $a_i = 0$, meaning not to abate.

		Player j	
		$a_j = 1$	$a_j = 0$
Player i	$a_i = 1$	$\gamma u(2) - c - \kappa\alpha_i$, $\gamma u(2) - c - \kappa\alpha_j$	$\gamma u(1) - c - \kappa\alpha_i + \alpha_i$, $\gamma u(1) - \alpha_i$
	$a_i = 0$	$\gamma u(1) - \alpha_j$, $\gamma u(1) - c - \kappa\alpha_j + \alpha_j$	$\gamma u(0)$, $\gamma u(0)$

TABLE 1. Payoffs in Climate Game with Protectionist Policy

Consequently, players will face a prisoner's dilemma: no individual player perceives abatement as privately optimal, and thus, $G = 0$ emerges as the unique Nash equilibrium. In contrast, the socially efficient outcome would be $G = 2$. Due to misaligned private and social incentives, the Nash equilibrium results in inefficiency. The following Lemma summarizes these insights.

LEMMA 1 (Nash equilibrium of pure climate game). *In the absence of protectionism ($\alpha_i = \alpha_j = 0$), the unique Nash equilibrium of the climate game is characterized by $a_i = a_j = 0$, resulting in $G = 0$. This equilibrium is Pareto-dominated by $a_i = a_j = 1$, yielding $G = 2$.*

2.2. BTN climate game with unregulated protectionism and known κ

As another benchmark, consider the game in which players can freely choose their level of protectionism α_i and in which the BTN cost parameter κ is known.

The payoff function in eq. (3) is such that the marginal cost of BTN policies is given by κ, while the transfer function in eq. (2) indicates that the marginal benefit of BTN policies is equal to one. The linearity of the transfer function implies that if a player opts to abate, it will do so with fully protectionist climate policies (i.e., $\alpha = 1$). The game structure implies that the payoff to each player depends on whether the other player abates, and the BTN cost parameter κ determines the relative costs of protectionist abatement.

First, consider the case when $\kappa < \underline{\kappa} \equiv 1 + \gamma\left[u(1) - u(0)\right] - c$. In this case, the marginal cost of protectionist abatement, given by κ, is less than the threshold marginal benefit $\underline{\kappa}$. Consequently, the payoff to player i is higher if it abates with full protectionism regardless of the other player's action. Therefore, protectionist abatement becomes a dominant strategy for both players, and the unique Nash equilibrium is $(a_i = 1, a_j = 1)$.

Next, for $\kappa > \overline{\kappa} \equiv 1 + \gamma\left[u(2) - u(1)\right] - c$, the marginal cost of protectionist abatement exceeds the threshold marginal benefit $\overline{\kappa}$. As a result, no abatement becomes a dominant strategy for both players, leading to the unique Nash equilibrium $(a_i = 0, a_j = 0)$.

Finally, for intermediate values of $\kappa \in (\underline{\kappa}, \overline{\kappa})$, protectionist abatement is beneficial if the other player also abates, but not if the other player does not abate. This creates a coordination problem in which both players prefer to match the other's action, leading to two possible Nash equilibria in pure strategies: $(a_i = 1, a_j = 1)$ and $(a_i = 0, a_j = 0)$.

Figure 1 shows equilibrium play as a function of the BTN cost parameter κ. The interval $(\underline{\kappa}, \overline{\kappa})$ presents a *window of multiplicity*, in which the BTN climate game becomes a stag hunt, a class of games of strategic complementarity with multiple equilibria.[5]

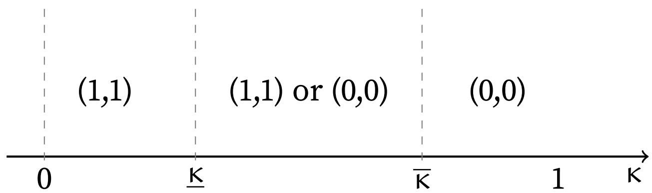

FIGURE 1. BTN Climate Game with Perfect Information – Equilibrium Play

Similarly, the welfare of each equilibrium outcome depends on the social payoff function, and the welfare-optimal outcome changes as a function of the BTN cost parameter κ.

If $\kappa < \kappa^w \equiv \gamma \left[u(2) - u(0)\right] - c$, the social benefit of both players abating exceeds the cost of protectionist climate policies. Therefore, the outcome $(a_i = 1, a_j = 1)$ maximizes welfare. In contrast, if $\kappa > \kappa^w$, the cost of abatement is too high, and welfare is maximized when neither player abates. Hence, the outcome $(a_i = 0, a_j = 0)$ is welfare-maximizing in this case. For intermediate values $\kappa \in (\underline{\kappa}^w, \overline{\kappa}^w)$, where $\underline{\kappa}^w = 2\gamma \left[u(1) - u(0)\right] - c$ and $\overline{\kappa}^w = 2\gamma \left[u(2) - u(1)\right] - c$, the outcome of only one player abating is welfare-minimizing. However, for $\kappa < \underline{\kappa}^w$, one player abating is preferable to no abatement. Similarly, for $\kappa > \overline{\kappa}^w$, one player abating is better than both players abating.

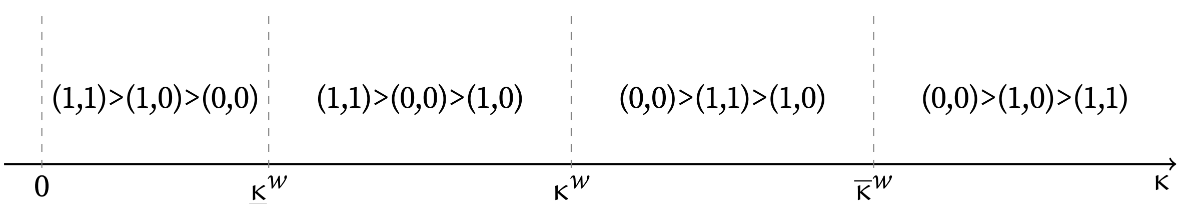

FIGURE 2. BTN Climate Game with Perfect Information – Welfare Ordering

Figure 2 illustrates the welfare ranking of different equilibrium outcomes of the unregulated BTN game as a function of the BTN cost parameter κ. It is never welfare-optimal for

[5]Fields and Lindequist (2024) empirically investigate the existence of climate policy risk spillovers from the US to the EU and find that financial markets expect the EU to align its climate policy stance with that of the US, providing suggestive evidence for the potential of multiple equilibria in global abatement efforts.

just one player to abate. However, an outcome in which only one player abates is minimizing welfare if $\kappa \in \left(\underline{\kappa}^w, \overline{\kappa}^w\right)$, i.e., a coordination failure is particularly welfare-damaging for intermediate values of κ.

The following Proposition summarizes the equilibrium and welfare effects of the unregulated BTN climate game with known BTN costs κ.

PROPOSITION 1 (Equilibrium and welfare effects of unregulated BTN climate game with known BTN costs). *Consider the game where players choose their level of protectionism, denoted by α, with known BTN cost parameter κ.*

- *If $\kappa < \underline{\kappa}$, both players abating is the only equilibrium. If, additionally, $\kappa < \kappa^w$, this equilibrium is socially efficient. However, if $\kappa^w < \kappa < \underline{\kappa}$, this equilibrium is inefficient.*
- *If $\kappa > \overline{\kappa}$, no player abating is the only equilibrium. If, additionally, $\kappa > \kappa^w$, this equilibrium is socially efficient. However, if $\kappa^w > \kappa > \overline{\kappa}$, this equilibrium is socially inefficient.*
- *If $\kappa \in [\underline{\kappa}, \overline{\kappa}]$, the welfare effects are ambiguous since the BTN climate game features multiple equilibria.*

When the marginal private cost of implementing protectionist climate policies κ is low, the private benefit each player gains from these policies exceeds the cost. As a result, both players find it individually rational to abate, leading to a socially efficient outcome if the total social benefit from abatement also outweighs the costs. In this case, the alignment between private and social incentives ensures that the equilibrium is optimal.

Conversely, when the marginal cost κ is high, the cost of protectionist policies outweighs the private benefit to each player. In this scenario, neither player chooses to abate, which can still be socially efficient if the overall cost of abatement exceeds the total social benefit. Here, private and social incentives again align, resulting in an optimal equilibrium.

However, when the marginal cost κ is at an intermediate level, the relationship between costs and benefits becomes more complex. In such cases, each player's decision depends on whether the other player abates, creating a coordination problem. This leads to multiple possible outcomes: either both players abate or neither does. The welfare implications are ambiguous because the players may fail to coordinate on the socially optimal outcome, resulting in inefficiency.

Thus, the equilibrium outcomes and their welfare effects in this game depend critically on the balance between the marginal cost of protectionist policies, the private benefits that individual players receive, and the overall social benefits of abatement. Misalignment between private and social incentives is most problematic in the intermediate-cost range, where coordination failures can prevent socially desirable outcomes.

In the next section, we show how the case of multiple equilibria can be resolved in the context of imperfect information about the BTN cost parameter κ. Using global games methodology, we derive the unique Nash equilibrium of the unregulated BTN climate game. As we will see, the threat of costly coordination failure remains relevant in this context and is crucial in understanding the welfare effects of protectionist climate policy.

3. Uncertainty about Beggar-Thy-Neighbor Costs

Consider the setup as before. Suppose that there is uncertainty about the cost associated with BTN policies, i.e., uncertainty about the cost parameter κ. This parameter is the same for both players. The 'true state of the world' is given by $z \in (-\infty, +\infty)$, which translates into the BTN cost parameter as follows:

$$\kappa = \zeta \exp(z), \tag{7}$$

where $\zeta > 0$ is a scale parameter. This functional form ensures that $\kappa \in (0, \infty)$. We assume that the true state of the world z is a random variable that is distributed with mean $\bar{z}$ and precision σ_z (i.e., variance $1/\sigma_z$): $z \sim \mathcal{N}(\bar{z}, 1/\sigma_z)$. This implies that the distribution of the BTN cost parameter κ is log-normal, i.e., its logarithm is normally distributed. Thus, the random variable κ has a probability density function given by

$$f(\kappa) = \frac{\sqrt{\sigma_z}}{\kappa\sqrt{2\pi}} \exp\left(-\frac{1}{2}\sigma_z \left(\ln(\kappa) - \ln(\zeta) - \bar{z}\right)^2\right). \tag{8}$$

This implies that the random variable $\ln(\kappa)$ is normally distributed with mean $\ln(\zeta) + \bar{z}$ and variance $1/\sigma_z$ (or precision σ_z). Furthermore, by the properties of the log-normal distribution, it holds that

$$\mathbb{E}(\kappa) = \zeta \exp\left(\bar{z} + \frac{1}{2\sigma_z}\right), \tag{9}$$

$$\mathrm{Var}(\kappa) = \zeta^2 \left(\exp\left(\frac{1}{\sigma_z}\right) - 1\right) \exp\left(2\bar{z} + \frac{1}{\sigma_z}\right). \tag{10}$$

These expressions provide a tractable map from beliefs about the true state of the world z to beliefs about the BTN cost parameter κ.

3.1. Common-belief game

Without any other sources of information about the true state of the world, both players hold the same belief about BTN costs with expected value $\mathbb{E}(\kappa) = \zeta \exp\left(\bar{z} + \frac{1}{2\sigma_z}\right)$. The equilibrium and welfare effects of this common-belief game immediately follow from the results in Proposition 1, which summarized the equilibrium and welfare effects of the unregulated BTN climate game with known BTN cost parameter κ. The key difference is that in the common-belief game, there is uncertainty regarding the true state of the world κ meaning that the welfare effects need to be evaluated in terms of (ex-ante) expected welfare.

COROLLARY 1 (Equilibrium and welfare effects of common-belief unregulated BTN climate game). *Consider the game where players choose their level of protectionism, denoted by α, with a common-belief of the expected BTN costs $\mathbb{E}(\kappa)$.*

- *If $\mathbb{E}(\kappa) < \underline{\kappa}$, both players abating is the only equilibrium. If, additionally, $\mathbb{E}(\kappa) < \kappa^w$, this equilibrium is ex-ante socially efficient. However, if $\kappa^w < \mathbb{E}(\kappa) < \underline{\kappa}$, this equilibrium is ex-ante inefficient.*
- *If $\mathbb{E}(\kappa) > \overline{\kappa}$, no player abating is the only equilibrium. If, additionally, $\mathbb{E}(\kappa) > \kappa^w$, this equilibrium is ex-ante socially efficient. However, if $\kappa^w > \mathbb{E}(\kappa) > \overline{\kappa}$, this equilibrium is ex-ante socially inefficient.*
- *If $\mathbb{E}(\kappa) \in [\underline{\kappa}, \overline{\kappa}]$, the welfare effects are ambiguous since the common-belief BTN climate game features multiple equilibria.*

Note that ex-post, it is possible that an equilibrium which was ex-ante socially efficient turns out to be inefficient. For instance, if $\mathbb{E}(\kappa) < \kappa^w$ (leading to both players optimally abating ex-ante), but the actual κ realized ex-post is significantly larger than $\mathbb{E}(\kappa)$, the cost of abatement outweighs the social benefits, rendering the equilibrium inefficient. However, because the players are risk-neutral—a result of the linearity of the payoff function—their decisions under uncertainty depend only on the expected value of BTN costs, $\mathbb{E}(\kappa)$. The variance or other higher moments of the belief distribution have no impact on ex-ante welfare considerations.

3.2. Beliefs and private information

Now suppose that there is uncertainty about the true state of the world z and that players observe a private signal about this state of the world. As before, z is a random variable with mean $\bar{z}$ and precision σ_z (i.e., variance $1/\sigma_z$): $z \sim \mathcal{N}(\bar{z}, 1/\sigma_z)$. Player i observes a private signal about z: $x_i = z + \epsilon_i$, where ϵ_i is normally distributed with mean 0 and precision σ_ϵ: $\epsilon_i \sim \mathcal{N}(0, 1/\sigma_\epsilon)$. Player i's signal is unobserved by player j. The signals are independent and identically distributed across players.

A strategy for a player is a rule of action that prescribes an action for each realization of the signal. A profile of strategies (one for each player) is an equilibrium if, conditional on the information available to player i and given the strategy followed by player j, the action prescribed by i's strategy maximizes her conditional expected utility. Treating the realization of i's signal as a possible 'type' of the player, we are solving for the Bayes Nash equilibrium of the imperfect-information game.

Player i uses the signal x_i to update her belief about the distribution of random variable z using Bayes rule. Suppose player i was sure that player j was going to follow a 'threshold' strategy where he abated only if the mean of his posterior belief of the true state of the world z was below $\hat{\rho}_j$, which means that

$$s_j(\rho_j) = \begin{cases} a_j = 0 \text{ if } \rho_j > \hat{\rho}_j, \\ a_j = 1 \text{ if } \rho_j < \hat{\rho}_j. \end{cases} \tag{11}$$

Under this strategy, player j will not abate if his posterior belief about the state z is sufficiently high so that his belief about the BTN cost parameter κ is sufficiently high as well (as a high z implies a high κ). The following Lemma summarizes player i's belief about the true state of the world as well as her belief about the other player's belief.

LEMMA 2 (Beliefs of players). *In the game with incomplete information, player i uses her private signal x_i to update her belief about the true state z using Bayes' rule. The posterior belief about z is distributed according to*

$$z|x_i \sim \mathcal{N}\left(\rho_i, \frac{1}{\sigma_z + \sigma_\epsilon}\right), \tag{12}$$

where $\rho_i \equiv \frac{\sigma_z \bar{z} + \sigma_\epsilon x_i}{\sigma_z + \sigma_\epsilon}$ is the mean of the player's posterior belief distribution. Player i's belief about player j's posterior belief ρ_j being less than some threshold $\hat{\rho}_j$ is given by:

$$q \equiv Prob(\rho_j < \hat{\rho}_j | \rho_i) = \Phi\left(\sqrt{\beta}\left(\hat{\rho}_j - \rho_i + \frac{\sigma_z}{\sigma_\epsilon}(\hat{\rho}_j - \bar{z})\right)\right), \tag{13}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution and $\beta \equiv \frac{\sigma_\epsilon(\sigma_z + \sigma_\epsilon)}{\sigma_z + 2\sigma_\epsilon}$ is the precision of the belief of player i about player j's signal given player i's signal.

Given these beliefs, player i's net expected gain from protectionist abatement is given by:

$$F \equiv \underbrace{q\gamma(u(2)-u(1)-c)+(1-q)\gamma(u(1)-u(0)-c)+\alpha_i}_{\text{expected benefit}} - \underbrace{\hat{\kappa}_i}_{\text{expected cost}}, \tag{14}$$

where q is the probability player j will abate. The expected benefit from protectionist abatement stems from an increase in the provision of the public good (which alone would not suffice to incentivize a player to abate), as well as an anticipated resource transfer from the other player resulting from protectionist policies, $\alpha_i = 1$. The cost of protectionist abatement, on the other hand, is captured by the player's posterior belief about the BTN costs, denoted by $\hat{\kappa}_i$.

Let $b_i(\hat{\rho}_j)$ denote player i's best response threshold strategy such that for all $\rho_i \leq b_i(\hat{\rho}_j)$, player i will abate (i.e., $F > 0$) and for all $\rho_i > b_i(\hat{\rho}_j)$ player i will not abate (i.e., $F < 0$). This best response cutoff strategy for player i is defined as the strategy that sets $F = 0$ at $\rho_i = b_i(\hat{\rho}_j)$ for a given threshold $\hat{\rho}_j$.

3.3. Existence and uniqueness of equilibrium

Note that player i is indifferent between abatement and no abatement if the net expected gain from abatement is equal to zero. Specifically,

$$\begin{aligned}
& F \equiv q\gamma(u(2)-u(1)) + (1-q)\gamma(u(1)-u(0)) + 1 - \left(c+\hat{\kappa}_i\right) = 0 \\
\iff \quad & \hat{\kappa}_i = 1+\gamma\left[q\left(u(2)-u(1)\right)+(1-q)\left(u(1)-u(0)\right)\right] - c \\
\iff \quad & \hat{\kappa}_i = \underline{\kappa} + q\left(\overline{\kappa}-\underline{\kappa}\right), \qquad (15)
\end{aligned}$$

where $\hat{\kappa}_i$ denotes the expected marginal BTN cost, and $\underline{\kappa} + q(\overline{\kappa}-\underline{\kappa})$ denotes the expected marginal BTN benefit. Given our analysis above, we know that

$$\overline{\kappa} = 1+\gamma\left[u(2)-u(1)\right]-c, \quad \underline{\kappa} = 1+\gamma\left[u(1)-u(0)\right]-c, \quad \hat{\kappa}_i = \mathbb{E}(\kappa|x_i) = \zeta\exp\left(\rho_i + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right),$$
$$q = \text{Prob}(\rho_j < \hat{\rho}_j|\rho_i) = \Phi\left(\sqrt{\beta}\left(\hat{\rho}_j - \rho_i + \frac{\sigma_z}{\sigma_\epsilon}(\hat{\rho}_j - \bar{z})\right)\right),$$

where $\beta = \frac{\sigma_\epsilon(\sigma_z+\sigma_\epsilon)}{\sigma_z+2\sigma_\epsilon}$ is the precision of player i's belief about player j's signal given player i's signal. As a result, the equilibrium threshold belief for player i, ρ_i^*, solves:

$$\zeta\exp\left(\rho_i^* + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right) = \underline{\kappa} + \Phi\left(\sqrt{\beta}\left(\hat{\rho}_j - \rho_i^* + \frac{\sigma_z}{\sigma_\epsilon}(\hat{\rho}_j - \bar{z})\right)\right)(\overline{\kappa}-\underline{\kappa}). \tag{16}$$

Similarly, player j's equilibrium threshold belief ρ_j^* is such that player j's marginal cost of abatement equals its marginal benefit. Assuming that players are identical and that the solution $\{\rho_i^*, \rho_j^*\}$ is unique, it is straightforward to show that the threshold beliefs will also be identical.

LEMMA 3. *Suppose players are identical, i.e.,* $\gamma_i = \gamma_j = \gamma$ *and* $c_i = c_j = c$. *Then, any unique equilibrium threshold beliefs* $\{\rho_i^*, \rho_j^*\}$ *are identical:* $\rho_i^* = \rho_j^* = \rho^*$.

Any unique equilibrium threshold belief ρ^* solves:

$$\zeta \exp\left(\rho^* + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right) = \underline{\kappa} + \Phi\left(\sqrt{\frac{\sigma_\epsilon(\sigma_z + \sigma_\epsilon)}{\sigma_z + 2\sigma_\epsilon}}\frac{\sigma_z}{\sigma_\epsilon}(\rho^* - \bar{z})\right)(\overline{\kappa} - \underline{\kappa})\,. \tag{17}$$

What remains to be shown is that such a unique threshold belief exists. The following Proposition establishes conditions such that the equilibrium threshold belief ρ^* exists and is unique.

PROPOSITION 2. *Suppose Assumption 1 holds. Then, if* $\theta \leq 2\pi\left(\frac{\underline{\kappa}}{\overline{\kappa}-\underline{\kappa}}\right)^2$ *there exists a unique equilibrium threshold belief* ρ^* *satisfying:*

$$\zeta \exp\left(\rho^* + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) = \underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\overline{\kappa} - \underline{\kappa}), \tag{18}$$

where $\theta = \beta\frac{\sigma_z^2}{\sigma_\epsilon^2}$ *and* $\beta = \frac{\sigma_\epsilon(\sigma_z+\sigma_\epsilon)}{\sigma_z+2\sigma_\epsilon}$. *In the resulting Bayesian Nash equilibrium, player i abates if and only if* $\rho_i < \rho^*$, *and does not abate otherwise.*

The Proposition demonstrates that as long as players' private signals are sufficiently precise, both players adopt a switching strategy centered around a critical threshold belief about BTN costs, denoted by ρ^*. At this threshold, the marginal cost of BTN policies, which arise from trade distortions, $MC \equiv \zeta \exp\left(\rho^* + \frac{1}{2(\sigma_\epsilon+\sigma_z)}\right)$, equals the marginal benefit, $MB \equiv \underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\overline{\kappa} - \underline{\kappa})$, which stems from the marginal utility of public good provision and the ability to extract resources from the other player through protectionist climate policies. The marginal benefit of BTN policies depends on whether the other player also engages in abatement, with the probability of abatement given by $q \equiv \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)$. Consequently, $\rho > \rho^*$ implies that $MC > MB$, which implies that a player finds it optimal not to abate, whereas if $\rho < \rho^*$, then $MC < MB$ and a player finds it optimal to abate. At the equilibrium threshold level ρ^*, the player is indifferent between abating and not abating.

To gain some more intuition for the uniqueness argument, consider Figure 3, which depicts the marginal benefit and marginal cost of protectionist climate policy for an

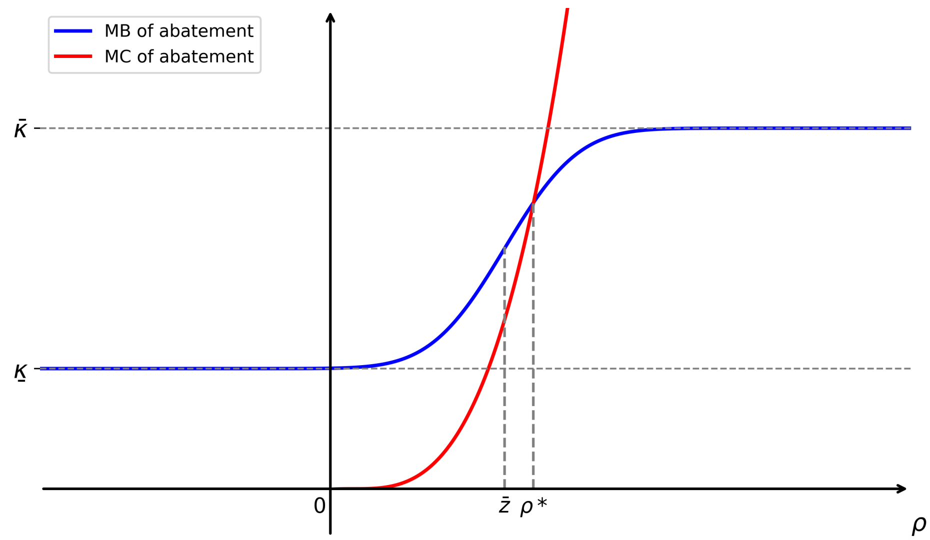

FIGURE 3. Existence and Uniqueness of Equilibrium

Notes: The figure illustrates the unique Bayesian Nash equilibrium. The marginal benefit (MB) of abatement is given by $\underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\overline{\kappa} - \underline{\kappa})$, and the marginal cost (MC) of abatement is given by $\zeta \exp\left(\rho^* + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right)$. The intersection of the marginal benefit of abatement and the marginal cost of abatement determines the equilibrium threshold belief ρ^*, which is such that a player abates for $\rho < \rho^*$ and does not abate for $\rho > \rho^*$.

individual player. The equilibrium threshold belief ρ^* is identified as the point where the marginal benefit curve intersects the marginal cost curve. Without further assumptions, multiple intersections between these curves are possible, which could lead to multiple solutions to eq. (18) and thus, multiple equilibria. To see this, note that the slope of the marginal benefit curve is governed by the slope of the standard normal CDF $\Phi(\cdot)$, which in turn depends on the value of θ, the relative precision of prior beliefs about BTN costs, σ_z, compared to the precision of private signals, σ_ϵ. Specifically, a low θ flattens the MB-curve. In the limit of $\theta \to 0$, the MB-curve becomes effectively a flat line at $\frac{1}{2}(\underline{\kappa} + \overline{\kappa})$, ruling out multiple intersections with the MC-curve. Conversely, as $\theta \to \infty$, the MB-curve effectively becomes a step function with a jump at $\rho = \bar{z}$. In this case, multiple intersections with the MC-curve become very likely, though not guaranteed. Therefore, a precise enough private signal, i.e., a low enough θ, is a sufficient (but not necessary) condition for the existence of a unique intersection and thus guarantees the existence of a unique equilibrium, as indicated in Proposition 2.

The term $\frac{\overline{\kappa}-\underline{\kappa}}{\underline{\kappa}}$ reflects the size of the 'window of multiplicity,' and equilibrium uniqueness requires this window to be sufficiently small. If $\underline{\kappa} = \overline{\kappa}$, the window collapses, and a unique equilibrium always exists, as either abating or not abating becomes a dominant strategy for both players. Importantly, the condition on $\frac{\overline{\kappa}-\underline{\kappa}}{\underline{\kappa}}$ is a sufficient, but not necessary, condition for uniqueness. Violations of this condition do not preclude uniqueness but indicate the possibility of multiple equilibria. If $\overline{\kappa}$ becomes very large, the marginal benefit of abatement becomes highly sensitive to players' beliefs about the other's actions, expanding the 'window of multiplicity' and amplifying the potential for multiple equilibria.

Extremes in the precision of information (θ) also shed light on equilibrium outcomes. For instance, if $\sigma_\epsilon = 0$, we revert to the common-belief game, where intermediate expected BTN costs reintroduce multiplicity. Conversely, if $\sigma_z = 0$, players rely solely on private signals, and the equilibrium cutoff corresponds to each player's probability of abatement being 1/2, ensuring uniqueness as the marginal benefit curve flattens.

Once the existence and uniqueness of a Bayesian Nash equilibrium in switching strategies are established, it is straightforward to conduct comparative statics of equilibrium abatement probability q^* with respect to the various parameters of the model. The following Corollary highlights the impact of a change in the prior belief about abatement costs $\bar{z}$ on equilibrium abatement probability.

COROLLARY 2. *The equilibrium abatement probability q^* is decreasing in $\bar{z}$.*

The intuition here is that as $\bar{z}$ increases, the expected cost of protectionist climate policy for any player increases. This reduces the incentives for either player to abate at the posterior belief ρ^*. As a result, both players lower their equilibrium cutoff ρ^*, which in turn implies that players play not abate for a larger set of signals, reducing their equilibrium abatement probability.

3.4. Expected welfare of protectionist climate policy

For the expected welfare of the BTN game, note that the true state of the world pins down all payoffs. More specifically, welfare for the four different outcomes of the game with symmetric players is given by:

$$W_{11} = 2\gamma u(2) - 2c - 2\kappa, \tag{19}$$

$$W_{10} = W_{01} = 2\gamma u(1) - c - \kappa, \tag{20}$$

$$W_{00} = 2\gamma u(0), \tag{21}$$

where $\kappa = \zeta \exp(z)$, i.e., the true BTN cost parameter depends on the realization of the state of the world z. Note that W_{00} is equal to the social welfare in the prisoner's dilemma of the pure climate game.

Now, the expected welfare of the climate game with protectionist policy is given by

$$EW = \int_{-\infty}^{\infty} \left[(q^*)^2 W_{11} + q(1-q^*)W_{10} + (1-q^*)qW_{01} + (1-q^*)^2 W_{00} \right] f(z)dz, \tag{22}$$

where $f(z) = \frac{\sqrt{\sigma_z}}{\sqrt{2\pi}} \exp\left(-\sigma_z \frac{(z-\bar{z})^2}{2}\right)$ is the pdf of the distribution of the state of the world z and q^* is the probability player i will abate given her private signal. The following Lemma shows that this integral can be calculated and simplified to a tractable expression in terms of the welfare thresholds derived in Section 2.

LEMMA 4 (Expected welfare of the unregulated BTN climate game with uncertainty about BTN costs). *The expected welfare of the unregulated BTN climate game with uncertainty about BTN costs is given by*

$$EW = W_{00} + 2q^* \left[q\kappa^w + (1-q^*)\underline{\kappa}^w - \mathbb{E}(\kappa) \right], \tag{23}$$

where $\kappa^w = \gamma\left[u(2) - u(0)\right] - c$, $\underline{\kappa}^w = 2\gamma\left[u(1) - u(0)\right] - c$, *and* $\mathbb{E}(\kappa) = \zeta\left(\bar{z} + \frac{1}{2\sigma_z}\right)$.

We see that if $\mathbb{E}(\kappa) < \underline{\kappa}^w$, then the BTN game will always be a welfare improvement over the pure climate game. Intuitively, when $\mathbb{E}(\kappa) < \underline{\kappa}^w$, then the expected welfare ordering is $W_{11} > W_{10} > W_{00}$. Thus, even if only one of the two players abates, such an equilibrium improves welfare compared to the pure climate game.

If $\underline{\kappa}^w < \mathbb{E}(\kappa) < \kappa^w$, we see that the welfare effects of the unregulated BTN game are ambiguous. If both players abate (which happens with probability $(q^*)^2$), then the BTN game is a welfare improvement over the pure climate game. However, if only one of the two players abates (which happens with probability $2q^*(1-q^*)$), then the BTN game leads to a reduction in welfare compared to the pure climate game. The overall effect of the BTN game on ex-ante welfare is thus ambiguous. The risk that only one of the two players ends up abating is the concern here from a welfare perspective. The lower the probability of that happening, the more likely it is that the unregulated BTN game is in fact a welfare improvement over the pure climate game.

If $\mathbb{E}(\kappa) > \kappa^w$, the unregulated BTN game is unambiguously reducing welfare compared to the pure climate game. In this case, the best possible outcome of the game is W_{00}, and no player should abate from a welfare perspective. The fact that there is a positive probability that one of the two players (or both) may end up abating reduces expected welfare compared to the prisoner's dilemma of the pure climate game.

4. Welfare Effects of Protectionist Policy

We are now in a position to ask: what are the welfare effects of protectionist climate policy? Consider the expected welfare from the unregulated BTN climate game with uncertain BTN costs which is given by

$$EW = W_{00} + 2q^*[q^* \kappa^w + (1 - q^*)\underline{\kappa}^w - \mathbb{E}(\kappa)]. \tag{24}$$

Comparison to social optimum. Let q^w be the probability of abatement that a social planner would impose to maximize the expected welfare of the two players. If the welfare gain from abating, given by κ^w, is larger than the expected cost of abating, given by $\mathbb{E}(\kappa)$, the social planner would want abatement to happen with certainty. Otherwise the social planner would not want abatement to happen at all. Specifically,

$$q^w = \begin{cases} 1 & \text{if } \mathbb{E}(\kappa) < \kappa^w, \\ 0 & \text{if } \mathbb{E}(\kappa) > \kappa^w. \end{cases} \tag{25}$$

As the equilibrium abatement probability q^* lies in the open interval $(0, 1)$ for any $\mathbb{E}(\kappa)$, players either under-abate (i.e., abate with a too low probability) or over-abate (i.e., abate with a too high probability). As the social planner's first best will always be either abatement or no abatement for certain, we know exactly when players are over-abating and under-abating depending on the value of κ^w in equilibrium. Specifically, the deviation from the first best is given by:

$$D(q^w, q^*) = \|q^w - q^*\| = \begin{cases} 1 - \Phi(\sqrt{\theta}(\rho^* - \bar{z})) & \text{if } \mathbb{E}(\kappa) < \kappa^w, \\ \Phi(\sqrt{\theta}(\rho^* - \bar{z})) & \text{if } \mathbb{E}(\kappa) > \kappa^w. \end{cases} \tag{26}$$

Comparison to pure climate game. Even though the unregulated BTN climate game may not achieve the social optimum, it may nonetheless represent an improvement over the prisoner's dilemma of the pure climate game. Note that we can rewrite the expected welfare of the unregulated BTN climate game in eq. (24) as

$$EW = W_{00} + 2q^*(\kappa^w - \underline{\kappa}^w)\left[q^* - \frac{\mathbb{E}(\kappa) - \underline{\kappa}^w}{\kappa^w - \underline{\kappa}^w}\right]. \tag{27}$$

Now, let q^P be the minimum probability of abatement required for the unregulated BTN climate game to be a welfare improvement over the pure climate game. Specifically,

(28)
$$q^P \equiv \begin{cases} 0 & \text{if } \mathbb{E}(\kappa) < \underline{\kappa}^w, \\ \frac{\mathbb{E}(\kappa)-\underline{\kappa}^w}{\kappa^w-\underline{\kappa}^w} & \text{if } \mathbb{E}(\kappa) \in [\underline{\kappa}^w, \kappa^w], \\ 1 & \text{if } \mathbb{E}(\kappa) > \kappa^w. \end{cases}$$

Therefore by construction, the unregulated BTN climate game is welfare improving the prisoner's dilemma if and only if equilibrium abatement probability is more than or equal to q^P. In other words,

(29)
$$EW - W_{00} \begin{cases} > 0 & \text{if } q^* > q^P, \\ = 0 & \text{if } q^* = q^P, \\ < 0 & \text{if } q^* < q^P. \end{cases}$$

As a result, the unregulated BTN climate game is a welfare improvement if the equilibrium abatement probability is sufficiently large.

4.1. The role of expected BTN costs

With regard to the effect of the prior mean $\bar{z}$ on optimal abatement probabilities, note that

(30)
$$q^w = \begin{cases} 1 & \text{if } \bar{z} < \bar{z}^w, \\ 0 & \text{if } \bar{z} > \bar{z}^w, \end{cases}$$

where $\bar{z}^w$ solves $\mathbb{E}(\kappa) = \kappa^w$. Further,

(31)
$$D(q^w, q^*) = \|q^w - q^*\| = \begin{cases} 1 - \Phi(\sqrt{\theta}(\rho^* - \bar{z})) & \text{if } \bar{z} < \bar{z}^w, \\ \Phi(\sqrt{\theta}(\rho^* - \bar{z})) & \text{if } \bar{z} > \bar{z}^w, \end{cases}$$

and

(32)
$$q^P = \begin{cases} 0 & \text{if } \bar{z} < \bar{z}', \\ \frac{\mathbb{E}(\kappa)-\underline{\kappa}^w}{\kappa^w-\underline{\kappa}^w} & \text{if } \bar{z} \in [\bar{z}', \bar{z}^w], \\ 1 & \text{if } \bar{z} > \bar{z}^w, \end{cases}$$

where $\bar{z}'$ solves $\mathbb{E}(\kappa) = \underline{\kappa}^w$ with $\mathbb{E}(\kappa) = \zeta \exp\left(\bar{z} + \frac{1}{2\sigma_z}\right)$ so that $\frac{\partial q^P}{\partial \bar{z}} \geq 0$.

Note that q^P is an increasing function of $\bar{z}$: when expected protectionism costs go up, the likelihood of abatement must also go up to ensure the gains of abatement are larger than the costs. Since q^* is a strictly decreasing function of $\bar{z}$ (see Corollary 2) and also bounded between 0 and 1, it must be the case that there exists some $\bar{z}''$ such that

$$q^* - q^P = \begin{cases} q^* - q^P > 0 & \text{if } \bar{z} < \bar{z}'', \\ q^* - q^P < 0 & \text{if } \bar{z} > \bar{z}''. \end{cases} \tag{33}$$

This $\bar{z}''$ sets the marginal social cost equal to the marginal social benefit of the unregulated BTN climate game. Specifically, $\bar{z}''$ solves the equation:

$$q^* \kappa^w + (1 - q^*)\underline{\kappa}^w = \zeta \exp\left(\bar{z}'' + \frac{1}{2\sigma_z}\right) \equiv \mathbb{E}(\kappa''). \tag{34}$$

The following Proposition summarizes the effects of expected BTN costs on the expected welfare of the unregulated BTN climate game.

PROPOSITION 3 (Welfare effects of protectionist climate policy). *Let $\bar{z}$ represent the prior mean belief about BTN costs in the unregulated BTN climate game. The equilibrium abatement probabilities and welfare outcomes behave as follows:*

1. *For low $\bar{z}$ ($\bar{z} < \bar{z}'$): The BTN climate game is a welfare improvement over the prisoner's dilemma of the pure climate game, even if only one player abates.*

2. *For intermediate-low $\bar{z}$ ($\bar{z}' < \bar{z} < \bar{z}''$): The BTN climate game is a welfare improvement over the prisoner's dilemma of the pure climate game only if both players abate. Equilibrium abatement probabilities are sufficiently large such that the BTN climate game leads to an expected welfare benefit.*

3. *For intermediate-high $\bar{z}$ ($\bar{z}'' < \bar{z} < \bar{z}^w$): The BTN climate game is a welfare improvement over the prisoner's dilemma of the pure climate game only if both players abate. Equilibrium abatement probabilities are so low that the BTN climate game leads to an expected welfare loss.*

4. *For high $\bar{z}$ ($\bar{z} > \bar{z}^w$): The BTN climate game can never be a welfare improvement over the prisoner's dilemma of the pure climate game. Equilibrium abatement probabilities are so high that the BTN climate game leads to an expected welfare loss.*

Figure 4 graphically illustrates the main insights from the Proposition. If $\bar{z} < \bar{z}^w$, abatement by both players is constrained efficient in the BTN climate game since $W_{11} > W_{00}$. If

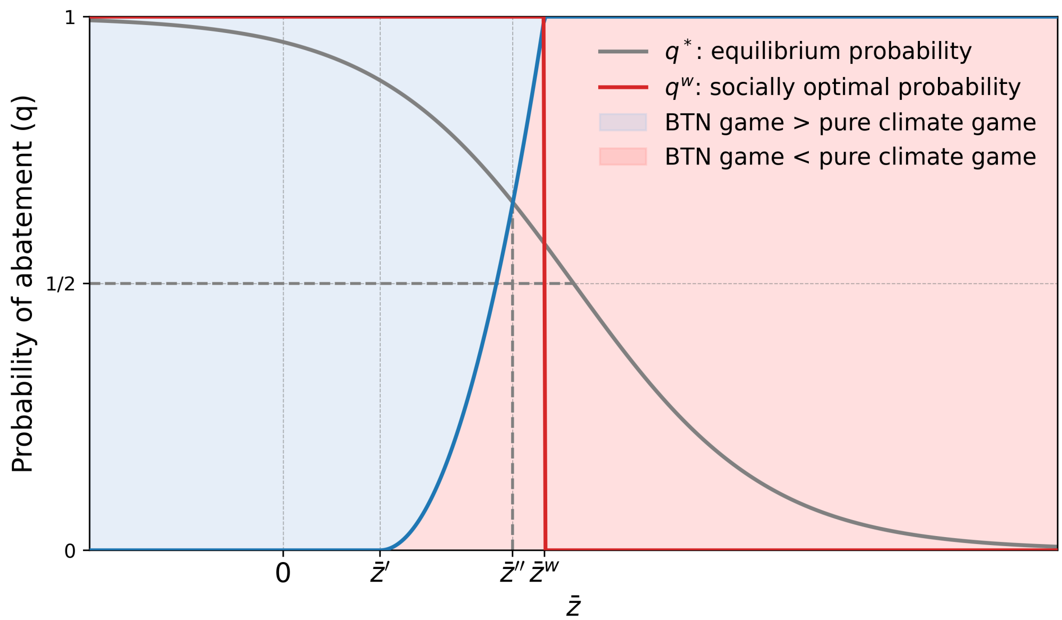

FIGURE 4. Welfare Effects of $\bar{z}$ (Mean of Prior Belief about BTN costs κ)

Notes: The figure illustrates the welfare effects of $\bar{z}$, the mean of prior beliefs about the costs associated with BTN policies. $\bar{z}'$ solves $\mathbb{E}(\kappa) = \underline{\kappa}^w$, $\bar{z}''$ solves $\mathbb{E}(\kappa) = q^*\kappa^w + (1-q^*)\underline{\kappa}^w$, and $\bar{z}^w$ solves $\mathbb{E}(\kappa) = \kappa^w$. The welfare rankings of payoffs are such that if $\bar{z} < \bar{z}'$ then $W_{11} > W_{10} > W_{00}$, if $\bar{z}' < \bar{z} < \bar{z}^w$ then $W_{11} > W_{00} > W_{10}$, and if $\bar{z} > \bar{z}^w$ then $W_{00} > W_{11}$. The blue line represents q^P, the minimum abatement probability for the BTN game to potentially be a welfare improvement over the pure climate game, and solves $q^P\kappa^w + (1-q^P)\underline{\kappa}^w = \mathbb{E}(\kappa)$.

$\bar{z} < \bar{z}' < \bar{z}^w$, the welfare ranking of the payoffs is $W_{11} > W_{10} > W_{00}$. This means that expected BTN costs are low enough for the BTN game to deliver a welfare benefit over the pure climate game (where no player abates), even if only one player abates. As a result, any probability of abatement in the BTN game will lead to a welfare improvement over the pure climate game.

For $\bar{z} \in (\bar{z}', \bar{z}^w)$, the welfare ranking shifts to $W_{11} > W_{00} > W_{10}$. Here, the expected BTN costs are higher, so the BTN climate game only generates a welfare benefit over the pure climate game if both players abate. Consequently, the equilibrium abatement probability must be large enough to offset the risk of only one player abating, which would otherwise result in welfare loss. As $\bar{z}$ increases, the welfare damage from unilateral abatement rises, requiring even larger equilibrium abatement probabilities for the BTN climate game to improve welfare over the pure climate game. Once $\bar{z} > \bar{z}''$, equilibrium abatement probabilities are no longer sufficient to compensate for the risk of one player abating, and the BTN game becomes welfare-damaging compared to the pure climate game.

Crucially, the distinction between the intermediate-low and intermediate-high regions of expected BTN costs lies in the impact of coordination failure (i.e., when only one player abates). In the intermediate-low region, the costs of such failure are low enough that the unregulated BTN climate game still results in a welfare improvement. However, in the intermediate-high region, the costs of coordination failure are significant enough to make the unregulated BTN climate game an expected welfare loss.

When $\bar{z}$ exceeds $\bar{z}^w$, abatement is no longer optimal in the BTN game (since $W_{00} > W_{11}$), and any non-zero equilibrium abatement probability will reduce welfare compared to the pure climate game. As equilibrium abatement probabilities approach zero for large $\bar{z}$, the expected payoff of the BTN climate game aligns with that of the pure climate game. The greatest disparity between socially optimal and equilibrium abatement probabilities occurs near $\bar{z}^w$, where socially optimal probabilities abruptly shift from 1 to 0, while equilibrium probabilities decrease gradually.

4.2. The role of regulation

Suppose that there is a regulating authority in place, such as the WTO, that can choose both countries' protectionist abatement level $\alpha \in [0, 1]$.

Equilibrium existence and uniqueness. Limiting protectionist policy to a level α changes the incentives for abatement for both countries. Specifically, for a given country j's threshold belief $\hat{\rho}_j$, country i's threshold belief, $\rho^*_{i\alpha}$, now solves the equation:

$$F_\alpha = \zeta \exp\left(\rho^*_{i\alpha} + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) - \left[\underline{\kappa}_\alpha + q_\alpha\left(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha\right)\right] = 0, \tag{35}$$

where $\overline{\kappa}_\alpha \equiv 1 + \frac{\gamma[u(2)-u(1)]-c}{\alpha}$, $\underline{\kappa}_\alpha \equiv 1 + \frac{\gamma[u(1)-u(0)]-c}{\alpha}$, and

$$q_\alpha \equiv \Phi\left(\sqrt{\beta}\left(\hat{\rho}_j - \rho^*_{i\alpha} + \frac{\sigma_z}{\sigma_\epsilon}(\hat{\rho}_j - \bar{z})\right)\right). \tag{36}$$

Note that if $\alpha < c - \gamma(u(1) - u(0))$, the regulation is such that countries would not abate even at zero costs to protectionism. Therefore, it is never optimal for regulators to set $\alpha < c - \gamma(u(1) - u(0))$. This gives us the following assumption for meaningful optimal regulatory policy:

Assumption 2. *Assume* $\alpha > c - \gamma(u(1) - u(0)) \equiv \alpha_{min}$.

As countries are symmetric, any equilibrium will be symmetric as well, i.e., $\rho^*_{i\alpha} = \rho^*_{j\alpha} = \rho^*_\alpha$.

Therefore, if an equilibrium exists, the equilibrium threshold ρ^*_α must solve

$$\zeta \exp\left(\rho^*_\alpha + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) = \underline{\kappa}_\alpha + \Phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha)\,. \tag{37}$$

This gives us the following Proposition for the existence of a unique equilibrium at a regulated level of protectionism α.

PROPOSITION 4. *Suppose Assumptions 1 and 2 hold. Then, if* $\theta \leq 2\pi\left(\frac{\underline{\kappa}_\alpha}{\overline{\kappa}_\alpha - \underline{\kappa}_\alpha}\right)^2$ *there exists a unique equilibrium* ρ^*_α *where the unique equilibrium,* ρ^*_α*, satisfies,*

$$\zeta \exp\left(\rho^*_\alpha + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) = \underline{\kappa}_\alpha + q^*_\alpha(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha), \tag{38}$$

where $\overline{\kappa}_\alpha = 1 + \frac{\gamma\{u(2)-u(1)\}-c}{\alpha}$, $\underline{\kappa}_\alpha = 1 + \frac{\gamma\{u(1)-u(0)\}-c}{\alpha}$*, and* $q^*_\alpha = \Phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)$.

The proof is nearly identical to the one for the case without regulation. If $\alpha = 1$, then we get back our Proposition 2 for unregulated protectionism.

Note that Proposition 4 implies that the equilibrium uniqueness breaks once the regulator imposes regulation that is too tight, i.e., an α that is too low.

COROLLARY 3. *The equilibrium of the regulated BTN climate game is unique as long as* $\alpha > \alpha_0 \equiv \alpha_{min} + \gamma(u(2) + u(0) - 2u(1))\sqrt{\frac{\theta}{2\pi}}$.

We see that α_0 declines as θ decreases. Therefore, for very precise signals it holds that $\alpha_0 \approx \alpha_{min}$. Note that since $\alpha_0 > \alpha_{min}$, at α_0, the equilibrium threshold ρ^*_0 elicits a positive equilibrium probability of abatement which we denote by q^*_0. If $\alpha \in (0, \alpha_0]$, the regulated BTN game can have multiple equilibria, which renders welfare analysis and comparative statics impossible. As a result, we will focus on the case in which $\alpha = \{0\} \cup (\alpha_0, 1]$. Note that the BTN game with $\alpha = 0$ returns the prisoner's dilemma of the pure climate game.

It is straightforward to show that the equilibrium abatement probability q^* increases in the level of protectionism α.

COROLLARY 4. *Suppose that* $\alpha > \alpha_0$*. Equilibrium abatement probability* q^* *is increasing in* α.

Intuitively, as α increases, the player's perceived benefit from abating increases. This implies that players will find it optimal to play abate for a larger set of posterior beliefs about the BTN cost parameter κ, which increases their belief thresholds and thus increases the probability that they will play abate.

Welfare effects of regulation. The expected welfare of the regulated climate game is given by:

$$EW = 2\gamma u(0) + 2q^*_\alpha[q^*_\alpha \kappa^w + (1 - q^*_\alpha)\underline{\kappa}^w - \alpha\mathbb{E}(\kappa)]. \tag{39}$$

Once again, let q^P_α be the minimum probability of abatement required for protectionist abatement level α to be welfare improving. Then we have,

$$q^P_\alpha = \begin{cases} 0 & \text{if } \alpha < \alpha', \\ \frac{\alpha\mathbb{E}(\kappa) - \underline{\kappa}^w}{\kappa^w - \underline{\kappa}^w} & \text{if } \alpha \in [\alpha', \alpha^w], \\ 1 & \text{if } \alpha > \alpha^w, \end{cases} \tag{40}$$

where $\alpha' = \frac{\underline{\kappa}^w}{\mathbb{E}(\kappa)}$ and $\alpha^w = \frac{\kappa^w}{\mathbb{E}(\kappa)}$. If $q^* > q^P_\alpha$ then protectionist abatement level α is welfare improving the pure climate game. On the other hand, if $q^* < q^P_\alpha$ then the pure climate game has greater social welfare than setting protectionist policy at α.

Further, the welfare impact of protectionism α is thus given by:

$$\frac{dEW}{d\alpha} = -2q^*_\alpha\mathbb{E}(\kappa) + 2\frac{dq^*_\alpha}{d\alpha}\left[\underline{\kappa}^w + 2q^*_\alpha\left(\kappa^w - \underline{\kappa}^w\right) - \alpha\mathbb{E}(\kappa)\right]. \tag{41}$$

The first term represents the marginal cost of protectionist climate policy: for a given probability of abatement, more protectionism increases the expected distortions from BTN policies (the 'BTN waste'). The second term represents the potential marginal benefit of protectionist climate policies: an increase in protectionism increases the equilibrium abatement probability (since $dq^*_\alpha/d\alpha > 0$ as per Corollary 4), which increases the expected welfare as long as the equilibrium success probability q^*_α is sufficiently large.

We know that an increase in equilibrium abatement probability q^*_α as a result of an increase in protectionism α has a positive welfare effect if and only if $q^*_\alpha > q^P_\alpha$. As a result, if $q^*_\alpha \leq q^P_\alpha$, a decrease in α (i.e., more regulation) will unambiguously increase expected welfare in the regulated BTN climate game (as $dEW/d\alpha < 0$).

These intuitions directly leads to the following Proposition:

PROPOSITION 5. *Consider the regulated BTN climate game with regulation* $\alpha \in [0, 1]$.

(a) *If* $\mathbb{E}(\kappa) < \frac{\underline{\kappa}^w}{\alpha_0}$, *then optimal protectionist policy* α^* *is positive.*

(b) *If* $\alpha_{min}\mathbb{E}(\kappa) > \kappa^w$ *then* $\alpha = 0$ *is welfare maximizing.*

(c) *If* $\mathbb{E}(\kappa) > \kappa^w$ *then there exists protectionist regulation* $\hat{\alpha} \in (\alpha_0, 1)$ *that is a welfare improvement over the unregulated protectionist climate game.*

We see that if expected BTN distortions are sufficiently small, then the regulator should allow for a positive level of protectionist climate policy. Intuitively, if at the lowest level of protectionist policies that generates a unique equilibrium, α_0, the equilibrium abatement probability q^*_α is larger than the required probability for the BTN game to welfare dominate the pure climate game, q^P_α, then we know that there is a region of values for α such that the BTN game welfare dominates the pure climate game.[6]

If expected distortions are sufficiently large, then the regulator wants to prevent countries from abating at all, which she can achieve by setting $\alpha = 0$, which reverts the game into the prisoner's dilemma of the pure climate game. Furthermore, if $\alpha = 1$, which is the level of protectionism in the unregulated BTN climate game, and it is the case that $q^*_1 < q^P_\alpha$, then the regulator will be able to improve welfare by reducing the level of protectionism α below 1.

4.3. The role of public information

In reality, there is no effective regulator. What the WTO can do instead of directly regulating protectionist policy is to provide public information about the welfare costs of protectionist policies. In terms of our model, we can conceptualize an increase in public information about distortions from protectionist policies as an increase in the precision σ_z of the prior belief about the true state of the world. For the remainder of this subsection, suppose that the regulator cannot regulate the level of protectionism α but can control the precision of the prior belief about BTN costs, σ_z.

Common-belief game. First, consider again the unregulated BTN climate game without private signals as a benchmark. In this game, both players hold the same prior belief about BTN costs with expected value $\mathbb{E}(\kappa) = \zeta \exp\left(\bar{z} + \frac{1}{2\sigma_z}\right)$. To simplify the analysis, we assume that the expected BTN costs are sufficiently low such that both players will play abate as the dominant strategy.

ASSUMPTION 3. *Assume that* $\mathbb{E}(\kappa) < \underline{\kappa}$.

Given this assumption, the unregulated BTN climate game without private signals generates a unique equilibrium in which both players abate.[7] If expected BTN costs are sufficiently low, then abatement will be the dominant strategy for both players. Note that in this common-belief game, abatement with full protectionism being a dominant strategy

[6] Note that this is a sufficiency condition only. It may be the case that optimal protectionist policy is positive even if this condition does not hold.

[7] See Corollary 1 for explanation.

is not necessarily socially optimal. We know from Corollary 1 that an equilibrium in which both players abate as a dominant strategy is socially inefficient if $\kappa^w < \underline{\kappa}$ and $\mathbb{E}(\kappa) \in \left(\kappa^w, \underline{\kappa}\right)$. That is, abatement can be socially inefficient if the joint BTN costs from protectionist policies are sufficiently large to outsize the positive welfare effects of abatement.

Private-information game. Now consider the BTN climate game with private signals. In this game, players are no longer sure about each other's beliefs as they do not observe each other's signals about the state of the world. As a result, players may have different posteriors about the state of the world. The unique Bayesian Nash equilibrium, in this case, consists of switching strategies, as shown in Proposition 2. If player i observes a high private signal about the state of the world, then she will believe that the true BTN costs are rather large, which makes abatement rather unattractive. At the same time, player i knows that a high private signal increases the likelihood that player j receives a rather high signal as well and may decide not to abate, which further reduces the expected payoffs from abatement. As a result, abatement will not be a dominant strategy in the private information game, even if it is in the common-belief game.

In the unregulated BTN game with private signals, the threshold belief ρ^* will be such that the marginal benefit of abatement equals its marginal cost at that threshold. If abatement is a dominant strategy in the common-belief game, then the equilibrium threshold belief in the private-information game ρ^* is above the prior belief $\bar{z}$ and increases in the precision of the prior belief, σ_z.

LEMMA 5. *Suppose Assumption 1 and Assumption 3 are fulfilled. Then,* $\rho^* > \bar{z}$ *and the equilibrium abatement probability increases in the precision of the prior belief:* $\frac{\partial q^*}{\partial \sigma_z} > 0$.

Intuitively, the equilibrium threshold belief ρ^* must be above the prior belief $\bar{z}$ if abatement is a dominant strategy in the common-belief game since a player will have to become more pessimistic about the true BTN costs after observing its private signal (i.e., it must expect a higher cost level) to find it optimal not to abate. Specifically, if a player became more optimistic about the true BTN costs after observing the private signal (i.e., if $\rho_i < \bar{z}$), the marginal benefit of abatement would be larger than its marginal cost at that posterior belief. This means the cut-off posterior ρ^*, which sets a player's expected marginal belief equal to her expected marginal cost equal, must be greater than $\bar{z}$.

The effect of prior precision σ_z on equilibrium abatement probability q^* is driven by two effects. First, an increase in the prior belief precision σ_z directly reduces the belief about expected BTN costs for any given signal x_i, i.e., it reduces $\hat{\kappa}_i = \mathbb{E}(\kappa|x_i) = \zeta \exp\left(\rho + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right)$, which increases player i's marginal benefit from abatement and which tends to increase abatement probability. Second, an increase in σ_z decreases the relative weight each player puts on its private signal x_i. As a result, player i becomes more certain about player j's posterior belief as the role of private signals is diminished in forming posterior beliefs. This implies that the precision of player i's belief that $\rho_j < \rho^*$, in which case player j abates, increases. Consequently, the probability that $\rho_j < \rho^*$ increases as long as $\rho^* > \bar{z}$, which is the case if abatement is a dominant strategy in the common-belief game.

We are now in a position to assess the welfare effects of changes in the public information precision σ_z.

PROPOSITION 6. *Assume Assumption 1 and Assumption 3 hold. Consider an increase in the precision of the public signal, σ_z, from σ_z to σ_z'. Let $\bar{\sigma}_z$ solve $\theta = 2\pi\left(\frac{\kappa}{\bar{\kappa}-\underline{\kappa}}\right)^2$ and assume that $\sigma_z < \bar{\sigma}_z$.*

(a) *Suppose that $\sigma_z' \leq \bar{\sigma}_z$. The welfare benefit of the unregulated BTN climate game relative to the prisoner's dilemma of the pure climate game increases if $\mathbb{E}(\kappa) < \kappa^w$, and decreases if $\mathbb{E}(\kappa) > \kappa^w$.*

(b) *Suppose that $\sigma_z' > \bar{\sigma}_z$. The equilibrium of the unregulated BTN climate game may no longer be unique in the private information game, and the welfare effects are ambiguous.*

The Proposition illustrates that more precise information provided by a regulator is not always improving the relative welfare of the BTN climate game. Specifically, when expected BTN costs are sufficiently high such that a social planner would want both players not to abate (in which case the pure climate game welfare dominates the unregulated BTN climate game), then more precise public information about BTN costs moves the equilibrium further away from the social optimum as it increases abatement probabilities that are already too high. If, however, expected BTN costs are low enough such that a social planner would want both players to abate, then an increase in the precision of public information moves equilibrium abatement probabilities in a welfare-improving direction.

We conclude that if expected BTN costs are sufficiently high, then it may be the case that a regulator will not improve welfare by providing more precise information about these high costs. Rather the opposite, a regulator would improve welfare by reducing the precision of the public information about expected BTN costs. Intuitively, more precise prior information makes players more likely to abate, which is not desirable when BTN costs are already quite high.

The second part of Proposition 6 points out that if public information becomes too precise, then the assumption on the relative precision of the private signal, θ, in Proposition 2 will be violated, which implies that the equilibrium in the unregulated BTN climate game may no longer be unique, which in turn renders welfare analysis impossible. Thus, the ability of a regulator to predictably influence the equilibrium of the unregulated BTN climate game through the provision of public information is limited by the fact that more precise public information may re-introduce multiple equilibria which limits the regulators ability to predictably influence the outcome of the unregulated BTN climate game.

5. Conclusion

This paper explores the welfare effects of protectionist climate policies in the context of the green transition. While economists have traditionally cautioned against protectionism due to its distortive effects, we demonstrate that, in the case of climate policies, these distortions can have strategic value by encouraging coordination between countries on climate action. We emphasize the nuanced role that protectionist policies can play in the green transition, noting that the World Trade Organization's blanket restrictions on trade-distorting industrial policies may, in certain cases, undermine global welfare. In a second-best world, allowing limited protectionist measures could help overcome the coordination failures that often plague international climate agreements.

Our findings highlight the importance of tailoring international regulations to the specific characteristics of green industrial policies. When expected distortions are low, protectionist policies may enhance welfare by promoting climate cooperation. When expected protectionist costs are high, the welfare impact is less of a concern, as countries will avoid pursuing such policies due to their prohibitive costs. However, when distortions from protectionist green policies are expected to be moderate, the risks of coordination failure outweigh the potential benefits, and stricter regulatory oversight may be necessary. Going forward, policy frameworks should consider these trade-offs carefully to strike a balance between encouraging climate action and minimizing the negative consequences of protectionism.

References

Altenburg, Tilman, and Dani Rodrik. 2017. "Green industrial policy: Accelerating structural change towards wealthy green economies." *Green Industrial Policy*.

Andres, Pia. 2023. "Industrial policy and global public goods provision: rethinking the environmental trade agreement."

Attinasi, Maria-Grazia, Lukas Boeckelmann, and Baptiste Meunier. 2023. "The economic costs of supply chain decoupling." *The World Economy*.

Attinasi, MG, L Boeckelmann, and M Baptiste. 2023. "Unfriendly friends: Trade and relocation effects of the US Inflation Reduction Act.CEPR." CEPR.

Balistreri, Edward J, and Christine McDaniel. 2024. "Waging a Global Trade War Alone: The Cost of Blanket Tariffs on Friend and Foe."

Barrett, Scott. 1994. "Self-enforcing international environmental agreements." *Oxford Economic Papers* 46 (Supplement_1): 878–894.

Barrett, Scott. 2005. "The theory of international environmental agreements." *Handbook of Environmental Economics* 3: 1457–1516.

Barrett, Scott, and Astrid Dannenberg. 2012. "Climate negotiations under scientific uncertainty." *Proceedings of the National Academy of Sciences* 109 (43): 17372–17376.

Barrett, Scott, and Astrid Dannenberg. 2022. "The decision to link trade agreements to the supply of global public goods." *Journal of the Association of Environmental and Resource Economists* 9 (2): 273–305.

Bayramoglu, Basak, Michael Finus, and Jean-François Jacques. 2018. "Climate agreements in a mitigation-adaptation game." *Journal of Public Economics* 165: 101–113.

Bistline, John, Geoffrey Blanford, Maxwell Brown, Dallas Burtraw, Maya Domeshek, Jamil Farbes, Allen Fawcett, Anne Hamilton, Jesse Jenkins, Ryan Jones et al. 2023. "Emissions and energy impacts of the Inflation Reduction Act." *Science* 380 (6652): 1324–1327.

Bown, Chad P, and Dan Wang. 2024. "Semiconductors and Modern Industrial Policy." *Journal of Economic Perspectives* 38 (4): 81–110.

Canayaz, Mehmet, Isil Erel, and Umit G Gurun. 2024. "When Protectionism Kills Talent." *Fisher College of Business Working Paper* (2024-03): 007.

Carlsson, Hans, and Eric Van Damme. 1993. "Global games and equilibrium selection." *Econometrica*: 989–1018.

Dahlström, Petter, Hans Lööf, Fredrik Sjöholm, and Andreas Stephan. 2023. "The EU's Competitive Advantage in the" clean-energy Arms Race"." Research Institute of Industrial Economics.

Ederington, Josh. 2010. "Should Trade Agreements Include Environmental Policy?" *Review of Environmental Economics and Policy* 4 (1): 84–102.

Evenett, Simon J. 2019. "Protectionism, state discrimination, and international business since the onset of the Global Financial Crisis." *Journal of International Business Policy* 2: 9–36.

Fajeau, Maxime, Niklas Garnadt, Veronika Grimm, Sébastien Jean, Thilo Kroeger, Camille Landais, Ulrike M Malmendier, Christian Ochsner, Hélène Paris, Thomas Philippon et al. 2023. "The US Inflation Reduction Act: How the EU is affected and how it should react." *Voxeu columns*.

Fajgelbaum, Pablo D, Pinelopi K Goldberg, Patrick J Kennedy, and Amit K Khandelwal. 2020. "The return to protectionism." *The Quarterly Journal of Economics* 135 (1): 1–55.

Fajgelbaum, Pablo D, and Amit K Khandelwal. 2022. "The economic impacts of the US–China trade war." *Annual Review of Economics* 14 (1): 205–228.
Farrokhi, Farid, and Ahmad Lashkaripour. 2024. "Can trade policy mitigate climate change." *Unpublished Working Paper*.
Fields, Micah, and David Lindequist. 2024. "Global spillovers of US climate policy risk: Evidence from EU carbon emissions futures." *Energy Economics* 139: 107931.
Finus, Michael. 2000. "Game theory and international environmental co-operation: A survey with an application to the Kyoto-Protocol." Nota di Lavoro.
Fischer, Carolyn. 2017. "Environmental protection for sale: strategic green industrial policy and climate finance." *Environmental and Resource Economics* 66: 553–575.
Frankel, David M, Stephen Morris, and Ady Pauzner. 2003. "Equilibrium selection in global games with strategic complementarities." *Journal of Economic Theory* 108 (1): 1–44.
Goldberg, Pinelopi K, Réka Juhász, Nathan J Lane, Giulia Lo Forte, and Jeff Thurk. 2024. "Industrial Policy in the Global Semiconductor Sector."
Gros, Daniel, Philipp-Leo Mengel, and Giorgio Presidente. 2023. "The EU and the US inflation reduction ect: No rose without thorns."
Hambel, Christoph, Holger Kraft, and Eduardo S Schwartz. 2018. "The carbon abatement game." National Bureau of Economic Research.
Harstad, Bård. 2012. "Climate contracts: A game of emissions, investments, negotiations, and renegotiations." *Review of Economic Studies* 79 (4): 1527–1557.
Harstad, Bård. 2024. "Contingent Trade Agreements." National Bureau of Economic Research.
Heijmans, Roweno JRK. 2022. "The Global Climate Game."
Heijmans, Roweno JRK. 2023. "Unraveling Coordination Problems." *arXiv preprint arXiv:2307.08557*.
Heitzig, Jobst, Kai Lessmann, and Yong Zou. 2011. "Self-enforcing strategies to deter free-riding in the climate change mitigation game and other repeated public good games." *Proceedings of the National Academy of Sciences* 108 (38): 15739–15744.
Hritonenko, Natali, Victoria Hritonenko, and Yuri Yatsenko. 2020. "Games with adaptation and mitigation." *Games* 11 (4): 60.
Irwin, Douglas A. 2011. *Trade policy disaster: lessons from the 1930s*.: MIT press.
Jorge, José, and Joana Rocha. 2015. "A primer on global games applied to macroeconomics and finance." *Journal of Economic Surveys* 29 (5): 869–886.
Juhász, Réka, and Nathan Lane. 2024. "The Political Economy of Industrial Policy." *Journal of Economic Perspectives* 38 (4): 27–54.
Juhász, Réka, Nathan Lane, and Dani Rodrik. 2023. "The new economics of industrial policy." *Annual Review of Economics* 16.
Kleimann, David, Niclas Poitiers, André Sapir, Simone Tagliapietra, Nicolas Véron, Reinhilde Veugelers, and Jeromin Zettelmeyer. 2023. "Green tech race? The US Inflation Reduction Act and the EU Net Zero Industry Act." *The World Economy* 46 (12): 3420–3434.
Kolstad, Charles D. 2007. "Systematic uncertainty in self-enforcing international environmental agreements." *Journal of Environmental Economics and Management* 53 (1): 68–79.
Lipsey, Richard G, and Kelvin Lancaster. 1956. "The general theory of second best." *The Review of Economic Studies* 24 (1): 11–32.
Martimort, David, and Wilfried Sand-Zantman. 2016. "A mechanism design approach to climate-

change agreements." *Journal of the European Economic Association* 14 (3): 669–718.

Mathevet, Laurent. 2010. "A contraction principle for finite global games." *Economic Theory* 42: 539–563.

Mazzucato, Mariana, and Dani Rodrik. 2023. "Industrial policy with conditionalities: a taxonomy and sample cases."

Morris, Stephen. 2016. *Global Games*. 1–7. London: Palgrave Macmillan UK.

Morris, Stephen, and Hyun Song Shin. 2000. "Rethinking multiple equilibria in macroeconomic modeling." *NBER Macroeconomics Annual* 15: 139–161.

Morris, Stephen, and Hyun Song Shin. 2003. *Global Games: Theory and Applications*. 56–114, Econometric Society Monographs: Cambridge University Press.

Nordhaus, William. 2015. "Climate clubs: Overcoming free-riding in international climate policy." *American Economic Review* 105 (4): 1339–1370.

Nordhaus, William. 2021. "Dynamic climate clubs: On the effectiveness of incentives in global climate agreements." *Proceedings of the National Academy of Sciences* 118 (45): e2109988118.

Robinson, Sherman, and Karen Thierfelder. 2024. "US international trade policy: Scenarios of protectionism and trade wars." *Journal of Policy Modeling* 46 (4): 723–739.

Rodrik, Dani. 2014. "Green industrial policy." *Oxford Review of Economic Policy* 30 (3): 469–491.

Rodrik, Dani. 2020. "Putting global governance in its place." *The World Bank Research Observer* 35 (1): 1–18.

Rubio, Santiago J, and Alistair Ulph. 2007. "An infinite-horizon model of dynamic membership of international environmental agreements." *Journal of Environmental Economics and Management* 54 (3): 296–310.

Sattich, Thomas, and Stella Huang. 2023. "Industrial competition-who is winning the renewable energy race?" In *Handbook on the Geopolitics of the Energy Transition,* 158–182: Edward Elgar Publishing.

Ulph, Alistair. 2004. "Stable international environmental agreements with a stock pollutant, uncertainty and learning." *Journal of Risk and Uncertainty* 29: 53–73.

Wood, Peter John. 2011. "Climate change and game theory." *Annals of the New York Academy of Sciences* 1219 (1): 153–170.

de Zeeuw, Aart. 2008. "Dynamic effects on the stability of international environmental agreements." *Journal of Environmental Economics and Management* 55 (2): 163–174.

Zehaie, Ficre. 2009. "The timing and strategic role of self-protection." *Environmental and Resource Economics* 44: 337–350.

Appendix

Appendix A. Proofs and Derivations

A.1. Proof of Lemma 1

Without protectionism (i.e., $\alpha_i = \alpha_j = 0$), the game reduces to a pure climate game where each player chooses whether to abate ($a_i = 1$) or not ($a_i = 0$). The utility from abatement is $\gamma u(G)$, where $G = a_i + a_j$, and the cost is c if $a_i = 1$.

By Assumption 1, the private benefit of abatement when the other player abates, $\gamma[u(2) - u(1)]$, is less than the private cost c, making abatement always privately inefficient. Therefore, the unique Nash equilibrium is ($a_i = a_j = 0$), leading to $G = 0$.

However, when both players abate ($a_i = a_j = 1$), the total benefit from the public good, $2\gamma u(2)$, exceeds the total cost $2c$, making abatement socially efficient. Thus, the Nash equilibrium ($a_i = a_j = 0$) is Pareto-dominated by the socially efficient outcome ($a_i = a_j = 1$). □

A.2. Proof of Proposition 1

The proof is provided in the main text.

A.3. Proof of Lemma 2

Since both z and x_i are normally distributed, the mean of player i's posterior belief distribution about z upon observing signal x_i is given by:

$$\rho_i = \frac{\sigma_z \bar{z} + \sigma_\epsilon x_i}{\sigma_z + \sigma_\epsilon}. \tag{A1}$$

Specifically, conditional on signal x_i, state z is normal with mean ρ_i and precision $\sigma_z + \sigma_\epsilon$, i.e., $z|x_i \sim \mathcal{N}\left(\rho_i, \frac{1}{\sigma_z+\sigma_\epsilon}\right)$. The corresponding belief about the true BTN cost parameter κ is given by

$$\ln(\kappa)|x_i \sim \mathcal{N}\left(\rho_i, \frac{1}{\sigma_z + \sigma_\epsilon}\right), \tag{A2}$$

so that,

$$\hat{\kappa}_i = \mathbb{E}(\kappa|x_i) = \zeta \exp\left(\rho_i + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right), \tag{A3}$$

$$\text{and} \quad Var(\kappa|x_i) = \zeta^2 \left(\exp\left(\frac{1}{\sigma_z + \sigma_\epsilon}\right) - 1\right) \exp\left(2\rho_i + \frac{1}{\sigma_z + \sigma_\epsilon}\right). \tag{A4}$$

We then see that,

$$\frac{\partial \hat{\kappa}_i}{\partial \rho_i} = \zeta \exp\left(\rho_i + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right) > 0, \tag{A5}$$

$$\frac{\partial \hat{\kappa}_i}{\partial \sigma_z} = -\frac{\zeta \exp\left(\rho_i + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right)}{2(\sigma_\epsilon + \sigma_z)^2} < 0, \tag{A6}$$

$$\text{and} \quad \frac{\partial \hat{\kappa}_i}{\partial \sigma_\epsilon} = -\frac{\zeta \exp\left(\rho_i + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right)}{2(\sigma_\epsilon + \sigma_z)^2} < 0. \tag{A7}$$

Intuitively, a larger posterior belief about the state z (given by ρ_i) translates into a higher belief about the BTN cost parameter. In addition, higher precision of either the prior distribution (higher σ_z) or the signal distribution (higher σ_ϵ) leads to smaller posterior beliefs about the BTN cost parameter, i.e., reduced uncertainty very tractably reduces the mean of the posterior belief distribution about the BTN cost parameter κ.

Now we find the relationship between the player i's belief of player j's posterior. Since $x_j = z + \epsilon_j$, the distribution of x_j conditional on ρ_i is normal with mean ρ_i and precision,

$$\frac{1}{\frac{1}{\sigma_z+\sigma_\epsilon} + \frac{1}{\sigma_\epsilon}} = \frac{\sigma_\epsilon(\sigma_z + \sigma_\epsilon)}{\sigma_z + 2\sigma_\epsilon} \equiv \beta. \tag{A8}$$

In other words,

$$x_j|\rho_i \sim \mathcal{N}\left(\rho_i, \frac{1}{\beta}\right). \tag{A9}$$

Furthermore,

$$\rho_j = \frac{\sigma_z \bar{z} + \sigma_\epsilon x_j}{\sigma_z + \sigma_\epsilon}. \tag{A10}$$

Note that for agent i the probability that agent j has a mean belief ρ_j less than some cutoff value $\hat{\rho}_j$ is given by:

$$\begin{aligned}
q \equiv \text{Prob}(\rho_j < \hat{\rho}_j | \rho_i) &= \text{Prob}\left(\frac{\sigma_z \bar{z} + \sigma_\epsilon x_j}{\sigma_z + \sigma_\epsilon} < \hat{\rho}_j | \rho_i\right) \\
&= \text{Prob}\left(x_j < \hat{\rho}_j + \frac{\sigma_z}{\sigma_\epsilon}(\hat{\rho}_j - \bar{z}) | \rho_i\right) \\
&= \Phi\left(\sqrt{\beta}\left(\hat{\rho}_j + \frac{\sigma_z}{\sigma_\epsilon}(\hat{\rho}_j - \bar{z}) - \rho_i\right)\right), \qquad \text{(A11)}
\end{aligned}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. □

A.4. Proof of Lemma 3

Note that if $\gamma_i = \gamma_j = \gamma$ and $c_i = c_j = c$, then $\overline{\kappa}_i = \overline{\kappa}_j = \overline{\kappa}$ and $\underline{\kappa}_i = \underline{\kappa}_j = \underline{\kappa}$. We will now show that when $\overline{\kappa}_i = \overline{\kappa}_j = \overline{\kappa}$ and $\underline{\kappa}_i = \underline{\kappa}_j = \underline{\kappa}$, then the equilibrium threshold beliefs will be identical, i.e., $\rho_i^* = \rho_j^* = \rho^*$. Consider our system of equations for the two threshold beliefs:

$$\zeta \exp\left(\rho_i^* + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right) = \underline{\kappa} + \Phi\left(\sqrt{\beta}\left(\rho_j^* - \rho_i^* + \frac{\sigma_\epsilon}{\sigma_\epsilon}(\rho_j^* - \bar{z})\right)\right)(\overline{\kappa} - \underline{\kappa}), \qquad \text{(A12)}$$

$$\zeta \exp\left(\rho_j^* + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right) = \underline{\kappa} + \Phi\left(\sqrt{\beta}\left(\rho_i^* - \rho_j^* + \frac{\sigma_\epsilon}{\sigma_\epsilon}(\rho_i^* - \bar{z})\right)\right)(\overline{\kappa} - \underline{\kappa}). \qquad \text{(A13)}$$

Define

$$F(\rho_i, \rho_j) = \zeta \exp\left(\rho_i^* + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right) - \underline{\kappa} - \Phi\left(\sqrt{\beta}\left(\rho_j^* - \rho_i^* + \frac{\sigma_\epsilon}{\sigma_\epsilon}(\rho_j^* - \bar{z})\right)\right)(\overline{\kappa} - \underline{\kappa}) \qquad \text{(A14)}$$

$$\text{and } G(\rho_i, \rho_j) = \zeta \exp\left(\rho_j^* + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right) - \underline{\kappa} - \Phi\left(\sqrt{\beta}\left(\rho_i^* - \rho_j^* + \frac{\sigma_\epsilon}{\sigma_\epsilon}(\rho_i^* - \bar{z})\right)\right)(\overline{\kappa} - \underline{\kappa}). \qquad \text{(A15)}$$

The system of equations for the two threshold beliefs are solved when $F(\rho_i, \rho_j) = G(\rho_i, \rho_j) = 0$. Due to the symmetry of the functions we have that

$$F(\rho_i^*, \rho_j^*) = G(\rho_j^*, \rho_i^*) \quad \text{and} \quad G(\rho_i^*, \rho_j^*) = F(\rho_j^*, \rho_i^*). \qquad \text{(A16)}$$

Therefore, if (ρ_i^*, ρ_j^*) is a solution of the system of equations for the two threshold beliefs, then (ρ_j^*, ρ_i^*) is also a solution. This means that if there is a unique solution it must be the case that

$$(\rho_i^*, \rho_j^*) = (\rho_j^*, \rho_i^*). \qquad \text{(A17)}$$

This condition holds if and only if $\rho_i^* = \rho_j^*$ which implies that a unique solution must be symmetric. $\square$

A.5. Proof of Proposition 2

We want to prove that the following equation:

$$\zeta \exp\left(\rho^* + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) = \underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\bar{\kappa} - \underline{\kappa}) \tag{A18}$$

has a unique solution for ρ^*, using Banach's Fixed-Point Theorem.[8]

First, we rewrite the equation in the form of a fixed-point equation by isolating ρ^* on one side:

$$\rho^* = g(\rho^*), \tag{A19}$$

where

$$g(\rho^*) = \ln\left(\frac{\underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\bar{\kappa} - \underline{\kappa})}{\zeta}\right) - \frac{1}{2(\sigma_\epsilon + \sigma_z)}. \tag{A20}$$

We now need to show that the function $g(\rho^*)$ is a contraction mapping. That is, there exists a constant $0 < \delta < 1$ such that for all $\rho_i^*, \rho_j^* \in \mathbb{R}$,

$$|g(\rho_i^*) - g(\rho_j^*)| \leq \delta|\rho_i^* - \rho_j^*|. \tag{A21}$$

To demonstrate this, we compute the derivative of $g(\rho^*)$ with respect to ρ^*:

$$g'(\rho^*) = \frac{\sqrt{\theta}(\bar{\kappa} - \underline{\kappa})\phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)}{\underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\bar{\kappa} - \underline{\kappa})}, \tag{A22}$$

where $\phi(\cdot)$ is the probability density function (PDF) of the standard normal distribution. Note that the largest possible value of $g'(\rho^*)$ is bounded above:

$$\max g'(\rho^*) < \frac{\sqrt{\theta}(\bar{\kappa} - \underline{\kappa})\frac{1}{\sqrt{2\pi}}}{\underline{\kappa}}. \tag{A23}$$

[8] See Mathevet (2010) for a discussion of equilibrium uniqueness in a class of global games in which the joint best-response function is a contraction.

Therefore, the bound on $g'(.)$ is guaranteed to be less than 1 if

$$\theta < 2\pi \left(\frac{\overline{\kappa}}{\overline{\kappa} - \underline{\kappa}} \right)^2 . \tag{A24}$$

This parameter condition implies that $g(\rho^*)$ is a contraction mapping, satisfying the condition for Banach's Fixed-Point Theorem. Since $g(\rho^*)$ is a contraction mapping, and the real line (or any closed interval) is a complete metric space under the standard Euclidean metric, we can apply Banach's Fixed-Point Theorem. The theorem guarantees the existence of a unique fixed point ρ^* such that

$$\rho^* = g(\rho^*). \tag{A25}$$

Therefore, there exists a unique ρ^* that solves the original equation if $\theta < 2\pi \left(\frac{\overline{\kappa}}{\overline{\kappa} - \underline{\kappa}} \right)^2$.

The last step is to show that when there is a unique equilibrium in switching strategies, then there cannot be any other equilibrium. This argument follows the standard argument of iterated deletion of dominated strategies outlined in Morris and Shin (2000) and is omitted here. □

A.6. Proof of Corollary 2

Consider the equilibrium condition given in Proposition 2 by eq. (18):

$$\zeta \exp \left(\rho^* + \frac{1}{2(\sigma_\epsilon + \sigma_z)} \right) = \left[\underline{\kappa} + \Phi \left(\sqrt{\theta}(\rho^* - \bar{z}) \right) (\overline{\kappa} - \underline{\kappa}) \right].$$

Define F as

$$F \equiv \zeta \exp \left(\rho^* + \frac{1}{2(\sigma_\epsilon + \sigma_z)} \right) - \left[\underline{\kappa} + \Phi \left(\sqrt{\theta}(\rho^* - \bar{z}) \right) (\overline{\kappa} - \underline{\kappa}) \right]. \tag{A26}$$

Note that in equilibrium, it holds that $F = 0$. Therefore, we can use the implicit function theorem to do comparative statics. That is, for a parameter $\psi = \{\gamma, c, \bar{z}, \sigma_z, \sigma_\epsilon, u(2), u(1), u(0)\}$ of the game we have that

$$\frac{\partial \rho^*}{\partial \psi} = -\frac{\frac{\partial F}{\partial \psi}}{\frac{\partial F}{\partial \rho^*}}. \tag{A27}$$

Note that our sufficiency condition for unique equilibrium ensures that $\forall\rho$ we have

$$\frac{\partial F}{\partial \rho} = \zeta \exp\left(\rho + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) - \sqrt{\theta}\phi\left(\sqrt{\theta}(\rho - \bar{z})\right)(\overline{\kappa} - \underline{\kappa}) > 0. \tag{A28}$$

Therefore, the relationship between ρ^* and any parameter ψ fully depends on the sign of $-\frac{\partial F}{\partial \psi}$. Further note that since $q^* = \text{Prob}(\rho < \rho^*|\rho^*) = \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)$, it holds that

$$\frac{\partial q^*}{\partial \psi} = \phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)\frac{\partial\sqrt{\theta}(\rho^* - \bar{z})}{\partial \psi}. \tag{A29}$$

The effect of changes in the mean of the prior distribution about the BTN cost parameter, $\bar{z}$, is thus obtained as

$$\begin{aligned} -\frac{\partial F}{\partial \bar{z}} &= -\sqrt{\theta}\phi(\sqrt{\theta}(\rho^* - \bar{z}))(\overline{\kappa} - \underline{\kappa}) < 0 \\ \implies \frac{\partial \rho^*}{\partial \bar{z}} &< 0. \end{aligned} \tag{A30}$$

This implies that

$$\frac{\partial q^*}{\partial \bar{z}} = \phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)\left[\sqrt{\theta}\underbrace{\frac{\partial \rho^*}{\partial \bar{z}}}_{<0} - \sqrt{\theta}\right] < 0. \tag{A31}$$

Therefore, an increase in $\bar{z}$ decreases equilibrium abatement probability q^*. □

A.7. Proof of Lemma 4

Using $\kappa = \zeta\exp(z)$ and $W_{10} = W_{10}$, we can write this integral as

$$\begin{aligned} EW = \int_{-\infty}^{\infty} \Bigg[&\left(\Phi\left(\sqrt{\theta}\left(\rho^* - \bar{z}\right)\right)\right)^2 (2\gamma u(2) - 2c - 2\zeta\exp(z)) \\ &+ 2\Phi\left(\sqrt{\theta}\left(\rho^* - \bar{z}\right)\right)\left(1 - \Phi\left(\sqrt{\theta}\left(\rho^* - \bar{z}\right)\right)\right)(2\gamma u(1) - c - \zeta\exp(z)) \\ &+ \left(1 - \Phi\left(\sqrt{\theta}\left(\rho^* - \bar{z}\right)\right)\right)^2 2\gamma u(0)\Bigg]\frac{\sqrt{\sigma_z}}{\sqrt{2\pi}}\exp\left(-\frac{\sigma_z(z - \bar{z})^2}{2}\right)dz. \end{aligned} \tag{A32}$$

This integral can be computed as

$$EW = 2\gamma u(0) + 2\left(2\gamma(u(1) - u(0)) - c - \mathbb{E}(\kappa)\right)\Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right) + 2\gamma\left(u(2) - u(1) - (u(1) - u(0))\right)\left[\Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)\right]^2, \tag{A33}$$

where $\mathbb{E}(\kappa) = \zeta \exp\left(\frac{1}{2\sigma_z} + \bar{z}\right)$ is the unconditional expected value of κ, i.e., the value before observing any private signals. We can rewrite the equation above using $q = \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)$ to

$$EW = 2\gamma u(0) + 2q^*\left(2\gamma(u(1) - u(0)) - c - \mathbb{E}(\kappa)\right) + (q^*)^2 2\gamma\left(u(2) - u(1) - (u(1) - u(0))\right). \tag{A34}$$

Note that the expression for expected welfare can be written more compactly as

$$\begin{aligned} EW &= (q^*)^2\left[2\gamma u(2) - 2c - 2\mathbb{E}(\kappa)\right] + 2q^*(1-q^*)\left[2\gamma u(1) - c - \mathbb{E}(\kappa)\right] + (1-q^*)^2\left[2\gamma u(0)\right] \\ &= q^2\mathbb{E}(W_{11}) + 2q(1-q)\mathbb{E}(W_{10}) + (1-q)^2 W_{00} \\ &= W_{00} + 2q^*\left[q^*\overline{\kappa}^w + (1-q^*)\underline{\kappa}^w - \mathbb{E}(\kappa)\right]. \end{aligned} \tag{A35}$$

Clearly, for the unregulated game to be a welfare improvement to the prisoner's dilemma game then each player's contribution to marginal social benefit of abating, given by $q^*\overline{\kappa}^w + (1-q^*)\underline{\kappa}^w$ must be greater than the cost of protectionism given by $\mathbb{E}(\kappa)$. □

A.8. Proof of Proposition 3

The proof is provided in the main text. □

A.9. Proof of Proposition 4

The proof follows the proof of Proposition 2.

The notable difference between the sufficiency condition for uniqueness in the regulated case versus the unregulated case is that the constraint on θ is stricter in the regulated case $\alpha \in (\alpha_{min}, 1]$. This uniqueness condition puts a constraint on α given by,

$$(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha)\sqrt{\frac{\theta}{2\pi}} < \underline{\kappa}_\alpha \tag{A36}$$

$$\Leftrightarrow \gamma(u(2) + u(0) - 2u(1))\sqrt{\frac{\theta}{2\pi}} < \alpha - c + \gamma(u(1) - u(0)) \tag{A37}$$

$$\Leftrightarrow \alpha > \alpha_{min} + \gamma(u(2) + u(0) - 2u(1))\sqrt{\frac{\theta}{2\pi}} \equiv \alpha_0, \tag{A38}$$

where α_{min} is the minimum value for α ensuring that players find it optimal to abate if $\kappa = 0$ regardless of each other's strategies, and α_0 denotes the minimum value of α to ensure equilibrium uniqueness. □

A.10. Proof of Corollary 3

The uniqueness condition for regulated protectionism puts the following lower bound on α:

$$\theta \leq 2\pi \left(\frac{\underline{\kappa}_\alpha}{\overline{\kappa}_\alpha - \underline{\kappa}_\alpha} \right)^2 \tag{A39}$$

$$\Leftrightarrow (\overline{\kappa}_\alpha - \underline{\kappa}_\alpha)\sqrt{\frac{\theta}{2\pi}} \leq \underline{\kappa}_\alpha \tag{A40}$$

$$\Leftrightarrow \gamma(u(2) + u(0) - 2u(1))\sqrt{\frac{\theta}{2\pi}} \leq \alpha - c + \gamma(u(1) - u(0)) \tag{A41}$$

$$\Leftrightarrow \alpha_0 \equiv \alpha_{min} + \underbrace{\gamma(u(2) + u(0) - 2u(1))\sqrt{\frac{\theta}{2\pi}}}_{>0} \leq \alpha. \tag{A42}$$

Thus, equilibrium uniqueness requires that $\alpha > \alpha_0 \geq \alpha_{min}$. □

A.11. Proof of Corollary 4

Suppose that $\alpha \in (\alpha_0, 1]$. We know from Corollary 3 that we get a unique equilibrium in this global game. Consider the equilibrium condition given in Proposition 4 by eq. (38):

$$\zeta \exp\left(\rho^*_\alpha + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) = \left[\underline{\kappa}_\alpha + \Phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha)\right].$$

Redefine F_α as

$$F_\alpha \equiv \zeta \exp\left(\rho^*_\alpha + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) - \left[\underline{\kappa}_\alpha + \Phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha)\right]. \tag{A43}$$

Note that in equilibrium, it holds that $F_\alpha = 0$. Therefore, we can use implicit function theorem to do comparative statics. Specifically, we have

$$\frac{\partial \rho^*_\alpha}{\partial \alpha} = -\frac{\frac{\partial F_\alpha}{\partial \alpha}}{\frac{\partial F}{\partial \rho^*}}. \tag{A44}$$

Now, given our sufficiency condition for unique equilibrium, we have

$$\frac{\partial F_\alpha}{\partial \rho^*_\alpha} = \zeta \exp\left(\rho^*_\alpha + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) - \sqrt{\theta}\phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)(\overline{\kappa}_\alpha - \underline{\kappa}_\alpha) > 0. \tag{A45}$$

Further partially differentiating F_α with respect to α we get

$$\begin{aligned}
-\frac{\partial F_\alpha}{\partial \alpha} &= \left[\frac{\partial \underline{\kappa}_\alpha}{\partial \alpha} + \Phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)\left(\frac{\partial \overline{\kappa}_\alpha}{\partial \alpha} - \frac{\partial \underline{\kappa}_\alpha}{\partial \alpha}\right)\right] \\
&= \underbrace{-\frac{\gamma[u(1) - u(0)] - c}{\alpha^2}}_{=\frac{1-\underline{\kappa}_\alpha}{\alpha}} - \underbrace{\Phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)}_{\equiv q^*_\alpha} \underbrace{\left(\frac{\gamma(u(2) + u(0) - 2u(1))}{\alpha^2}\right)}_{=\frac{\overline{\kappa}_\alpha - \underline{\kappa}_\alpha}{\alpha}} \\
&= \frac{1 - \overbrace{\left(q^*_\alpha \overline{\kappa}_\alpha + \left(1 - q^*_\alpha\right)\underline{\kappa}_\alpha\right)}^{=\hat{\kappa}_\alpha}}{\alpha} > 0,
\end{aligned} \tag{A46}$$

where both $\underline{\kappa}_\alpha$ and $\overline{\kappa}_\alpha$ are less than one individually.

Combining these values into Equation (A44) we get

$$\frac{\partial \rho^*}{\partial \alpha} > 0. \tag{A47}$$

Further note that since $q^*_\alpha = \text{Prob}(\rho_\alpha < \rho^* | \rho^*) = \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)$, it holds that

$$\frac{\partial q^*_\alpha}{\partial \alpha} = \phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)\frac{\partial \sqrt{\theta}(\rho^* - \bar{z})}{\partial \alpha}. \tag{A48}$$

We thus get that

$$\frac{\partial q^*_\alpha}{\partial \alpha} = \phi\left(\sqrt{\theta}(\rho^*_\alpha - \bar{z})\right)\sqrt{\theta}\underbrace{\frac{\partial \rho^*_\alpha}{\partial \alpha}}_{>0} > 0, \tag{A49}$$

which implies that an increase in α increases equilibrium abatement probability q^*_α. □

A.12. Proof of Proposition 5

Part (a). For a positive α to maximize welfare, the expected welfare from the regulated BTN climate game should be higher than the expected welfare of the prisoner's dilemma in the pure climate game, which is given by $W_{00} = 2\gamma u(0)$. The expected welfare of the regulated BTN climate game is given by:

$$EW = W_{00} + 2q_\alpha^*[q_\alpha^*\kappa^w + (1-q^*)\underline{\kappa}^w - \alpha\mathbb{E}(\kappa)]. \tag{A50}$$

We obtain the net welfare gain of the regulated BTN climate game relative to the pure climate game as

$$EW - W_{00} = 2q_\alpha^*[q_\alpha^*\kappa^w + (1-q_\alpha^*)\underline{\kappa}^w - \alpha\mathbb{E}(\kappa)]. \tag{A51}$$

The net welfare gain is positive if $q_\alpha^* > 0$ and if

$$q_\alpha^*\kappa^w + (1-q_\alpha^*)\underline{\kappa}^w - \alpha\mathbb{E}(\kappa) > 0 \tag{A52}$$

$$\iff \alpha\mathbb{E}(\kappa) < q_\alpha^*\kappa^w + (1-q_\alpha^*)\underline{\kappa}^w \tag{A53}$$

Since the lower bound for α is α_0 for q_α^* to be positive and unique, a sufficient condition for the existence of a positive optimal α^* is given by

$$\mathbb{E}(\kappa) < \frac{q_\alpha\kappa^w + (1-q_\alpha)\underline{\kappa}^w}{\alpha_0}. \tag{A54}$$

If this condition is satisfied, then there exists an $\alpha > 0$ that improves welfare compared to the case of the pure climate game in which $\alpha = 0$.

Part (b). Suppose $\alpha_{min}\mathbb{E}(\kappa) > \kappa^w$. Then Equation (A51) is negative for all values of $\alpha > \alpha_{min}$ as $q_\alpha \in [0,1]$ and $\kappa^w > \underline{\kappa}^w$. Since abatement can only be induced if $\alpha > \alpha_{min}$, there is no regulation $\alpha > 0$ that welfare improves from the prisoners' dilemma game and $\alpha = 0$ is optimal regulation policy.

Part (c). Suppose $\mathbb{E}(\kappa) > \kappa^w$. Then, there exists some $\hat{\alpha} \in (\alpha_0, 1)$ such that $\mathbb{E}(\kappa) > \hat{\alpha}\mathbb{E}(\kappa) > \kappa^w > \underline{\kappa}^w$. Since $\hat{\alpha} < 1$, we know that $\hat{\alpha}\mathbb{E}(\kappa) < \mathbb{E}(\kappa)$ and using Corollary 4 we get that $q_{\hat{\alpha}}^* < q_1^*$. This implies

$$\begin{aligned} q_1^*(q_1^*\underbrace{[\kappa^w - \mathbb{E}(\kappa)]}_{<0} + (1-q_1^*)\underbrace{[\underline{\kappa}^w - \mathbb{E}(\kappa)]}_{<0}) &< q_{\hat{\alpha}}^*(q_{\hat{\alpha}}^*[\kappa^w - \mathbb{E}(\kappa)] + (1-q_{\hat{\alpha}}^*)[\underline{\kappa}^w - \mathbb{E}(\kappa)]) \\ &< q_{\hat{\alpha}}^*(q_{\hat{\alpha}}^*[\kappa^w - \hat{\alpha}\mathbb{E}(\kappa)] + (1-q_{\hat{\alpha}}^*)[\underline{\kappa}^w - \hat{\alpha}\mathbb{E}(\kappa)]). \qquad \text{(A55)} \end{aligned}$$

Therefore, using eq. (A51) we see that the expected welfare at $\hat{\alpha}$ is higher than that at $\alpha = 1$. $\square$

A.13. Proof of Lemma 5

We will first prove that Assumption 3 implies that $\rho^* > \bar{z}$ and then show that it implies that $\frac{\partial q^*}{\partial \sigma_z} > 0$.

Proof that $\rho^* > \bar{z}$. Note that $\mathbb{E}(\kappa) < \underline{\kappa}$ means that $\zeta \exp\left(\bar{z} + \frac{1}{2\sigma_z}\right) < \underline{\kappa}$. It holds that $\zeta \exp\left(\bar{z} + \frac{1}{2\sigma_z}\right) > \zeta \exp\left(\bar{z} + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right)$ since $\sigma_\epsilon > 0$. As a result, the assumption of $\mathbb{E}(\kappa) < \underline{\kappa}$ implies $\zeta \exp\left(\bar{z} + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right) < \underline{\kappa}$.

Now, note that at $\rho = \bar{z}$, it holds that $q = \frac{1}{2}$, so that the marginal benefit of abatement is given by $\frac{1}{2}(\underline{\kappa} + \overline{\kappa}) > \underline{\kappa}$, while the marginal cost is given by $\zeta \exp\left(\bar{z} + \frac{1}{2(\sigma_z+\sigma_\epsilon)}\right)$. As a result, at $\rho = \bar{z}$, the assumption of $\mathbb{E}(\kappa) < \underline{\kappa}$ implies that the marginal benefit of abatement is larger than the marginal cost, implying that the equilibrium threshold belief ρ^* must be larger than $\bar{z}$.

Proof that $\frac{\partial q^*}{\partial \sigma_z} > 0$. Defining F as

(A56) $$F \equiv \zeta \exp\left(\rho^* + \frac{1}{2(\sigma_\epsilon + \sigma_z)}\right) - \left[\underline{\kappa} + \Phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)(\overline{\kappa} - \underline{\kappa})\right],$$

we have that

(A57) $$\frac{\partial \rho^*}{\partial \sigma_z} = -\frac{\frac{\partial F}{\partial \sigma_z}}{\frac{\partial F}{\partial \rho^*}}.$$

We already know from the proof of Corollary 2 that $\frac{\partial F}{\partial \rho^*} > 0$. Further,

(A58)
$$-\frac{\partial F}{\partial \sigma_z} = \zeta \exp\left(\rho + \frac{1}{2(\sigma_z + \sigma_\epsilon)}\right)\frac{1}{2(\sigma_z + \sigma_\epsilon)^2} + \phi\left(\sqrt{\theta}(\rho - \bar{z})\right)\frac{1}{2}\theta^{-\frac{1}{2}}\underbrace{\frac{\partial \theta}{\partial \sigma_z}}_{>0}\underbrace{(\rho - \bar{z})}_{>0}(\overline{\kappa} - \underline{\kappa}) > 0,$$

where

(A59) $$\frac{\partial \theta}{\partial \sigma_z} = -\frac{\sigma_z^2(\sigma_\epsilon + \sigma_z)}{\sigma_\epsilon(2\sigma_\epsilon + \sigma_z)^2} + \frac{\sigma_z^2}{\sigma_\epsilon(2\sigma_\epsilon + \sigma_z)} + \frac{2\sigma_z(\sigma_\epsilon + \sigma_z)}{\sigma_\epsilon(2\sigma_\epsilon + \sigma_z)} > 0.$$

This shows that $\frac{\partial \rho^*}{\partial \sigma_z} > 0$ as long as $\rho^* > \bar{z}$, which is guaranteed by Assumption 3 as shown above. Now,

$$\frac{\partial q^*}{\partial \sigma_z} = \phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)\left(\frac{\partial \sqrt{\theta}(\rho^* - \bar{z})}{\partial \sigma_z}\right)$$

$$= \phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)\left(\sqrt{\theta}\frac{\partial \rho^*}{\partial \sigma_z} + (\rho^* - \bar{z})\frac{\partial \sqrt{\theta}}{\partial \sigma_z}\right)$$

$$= \phi\left(\sqrt{\theta}(\rho^* - \bar{z})\right)\left(\sqrt{\theta}\underbrace{\frac{\partial \rho^*}{\partial \sigma_z}}_{>0} + \frac{\theta^{-\frac{1}{2}}}{2}\underbrace{(\rho^* - \bar{z})}_{>0}\underbrace{\frac{\partial \theta}{\partial \sigma_z}}_{>0}\right). \tag{A60}$$

Thus we conclude that $\frac{\partial q^*}{\partial \sigma_z} > 0$ when $\mathbb{E}(\kappa) < \underline{\kappa}$. □

A.14. Proof of Proposition 6

Part (a). We know from Proposition 1 that both players abating improves expected welfare over the prisoner's dilemma of the pure climate game if $\mathbb{E}(\kappa) < \min\{\kappa^w, \underline{\kappa}\}$. Assumption 3 implies that $\mathbb{E}(\kappa) < \underline{\kappa}$. Thus, if $\kappa^w > \underline{\kappa}$, then we know that expected welfare increases as equilibrium abatement increases, and any increase in equilibrium abatement probability triggered by an increase in the precision of public information σ_z will increase expected welfare.

However, if $\kappa^w < \underline{\kappa}$ and $\mathbb{E}(\kappa) \in \left(\kappa^w, \underline{\kappa}\right)$, then both players abating is not socially efficient and will reduce expected welfare compared to the pure climate game. In this case, an increase in equilibrium abatement probability triggered by an increase in σ_z will reduce expected welfare.

Part (b). We know from Proposition 2 that a unique equilibrium is only guaranteed as long as $\theta \leq 2\pi\left(\frac{\underline{\kappa}}{\kappa - \underline{\kappa}}\right)^2$. Since $\frac{\partial \theta}{\partial \sigma_z} > 0$, there will be a threshold value for $\bar{\sigma}_z$ where $\theta = 2\pi\left(\frac{\underline{\kappa}}{\kappa - \underline{\kappa}}\right)^2$ and such that for $\sigma_z > \bar{\sigma}_z$, equilibrium uniqueness is no longer guaranteed. □