Appendix: Information and Climate (In)action

Contents

Appendix A: Formal Proofs A-1
Appendix B: Additional Descriptive Figures A-21
Appendix C: Exxon Source Documents A-23

A Formal Proofs

Proofs of Domestic Politics Model

Domestic Politics: Equilibrium

I prove Proposition 1 with a series of two lemmas. The first establishes equilibrium behavior in the climate policy subgame, and the second characterizes the optimal misreporting level given this equilibrium behavior.

Define $\hat{\sigma}(\theta, \omega)$ as the voter's belief that probability that the politician chooses $a = 1$ when she is of type θ and the state of the world is ω. Define $B(\theta, a, s)$ as proportional to the *ex ante* probability that a politician of type θ chooses action a and signal s is realized. Then we have

$$B(\theta, 1, s) = P(\theta)\Big(\pi\hat{\sigma}(\theta, 1) + (1-\pi)\hat{\sigma}(\theta, 0)\frac{P(s|s \neq \omega)}{P(s|s = \omega)}\Big).$$
$$B(\theta, 0, s) = P(\theta)\Big(\pi(1 - \hat{\sigma}(\theta, 1)) + (1-\pi)(1 - \hat{\sigma}(\theta, 0))\frac{P(s|s \neq \omega)}{P(s|s = \omega)}\Big).$$

This means that the voter's posterior belief that the politician is competent, following policy choice a, is given by

$$\mu(a, s) = \frac{P(a, s|\theta = 1)P(\theta = 1)}{P(a, s|\theta = 1)P(\theta = 1) + P(a, s|\theta = 0)P(\theta = 0)} = \frac{B(1, a, s)}{B(1, a, s) + B(0, a, s)}.$$

Lemma A.1 *A unique cutoff $\tilde{x}^*$ exists, admitting a unique perfect Bayesian equilibrium to the climate policy subgame. A politician of type θ chooses policy $a = 1$ given signal x^θ with probability $\sigma^*(\theta, x^\theta) \in [0, 1]$. These probabilities are*

$$\sigma^*(1, x^1) = x^1 = \omega.$$
$$\sigma^*(0, x^0) = 1 - G(\tilde{x}^*; \omega).$$

Upon observing policy a and signal s, the voter reelects the politician with probability $F(\mu^*(a, s; \tilde{x}^*))$.

Proof of Lemma A.1: It is straightforward that following any history in which the politician chooses policy a and the voter observes signal s the voter has posterior belief $P(\theta = 1|a, s) = \mu(a, s)$, the voter reelects the politician if and only if $\mu(a, s) \geq \varepsilon$, which occurs with probability $F(\mu(a, s))$.

The competent politician, whose signal of ω is perfect, always chooses $a = 1$ following signal $x^1 = 1$:

$$1+\beta F(\mu(1,0))+(1-\beta)F(\mu(1,1)) \geq \beta F(\mu(0,0))+(1-\beta)F(\mu(0,1)) \Leftrightarrow 1 \geq -\beta\Delta(0)-(1-\beta)\Delta(1).$$

Similarly, she never chooses $a = 1$ following signal $x^1 = 0$:

$$F(\mu(1,0)) \leq 1 + F(\mu(0,0)) \;\Leftrightarrow\; \Delta(0) \leq 1.$$

Given the value of her private signal $x^0 = x$, the incompetent politician's posterior belief about the state is $\eta(x) = P(\omega = 1|x) = \frac{\pi g(x;1)}{\pi g(x;1)+(1-\pi)g(x;0)}$. Write $\Delta(s) = F(\mu(1,s)) - F(\mu(0,s))$. The incompetent type therefore chooses $a = 1$ if and only if

$$\begin{aligned}
&\eta(x) + \beta\eta(x)F(\mu(1,0)) + (1-\beta)\eta(x)F(\mu(1,1)) + (1-\eta(x))F(\mu(1,0)) \geq \\
&\qquad (1-\eta(x)) + \beta\eta(x)F(\mu(0,0)) + (1-\beta)\eta(x)F(\mu(0,1)) + (1-\eta(x))F(\mu(0,0)) \\
&\quad \Leftrightarrow\; \eta(x) \geq \frac{1-\Delta(0)}{2+(1-\beta)(\Delta(1)-\Delta(0))}.
\end{aligned}$$

Define $\tilde{x}$ as the signal that solves

$$2\eta(\tilde{x}) - 1 + (1-\beta)\eta(\tilde{x})\Delta(1;\tilde{x}) + (1-\eta(\tilde{x}) + \beta\eta(\tilde{x}))\Delta(0;\tilde{x}) = 0, \tag{1}$$

where the cutoff $\tilde{x}$ induces voter posterior beliefs

$$\begin{aligned}
\mu^*(1,0;\tilde{x}) &= \frac{\tau\pi}{\tau\pi + (1-\tau)\pi(1-G(\tilde{x};1)) + (1-\tau)(1-\pi)\frac{\beta}{1-\beta}(1-G(\tilde{x};0))}. \\
\mu^*(1,1;\tilde{x}) &= \frac{\tau}{\tau + (1-\tau)(1-G(\tilde{x};1))}. \\
\mu^*(0,0;\tilde{x}) &= \frac{\tau(1-\pi)\frac{\beta}{1-\beta}}{\tau(1-\pi)\frac{\beta}{1-\beta} + (1-\tau)(\pi G(\tilde{x};1) + (1-\pi)\frac{\beta}{1-\beta}G(\tilde{x};0))}. \\
\mu^*(0,1;\tilde{x}) &= 0.
\end{aligned}$$

Differentiating Equation 1 with respect to $\tilde{x}$ yields

$$2\frac{\partial\eta(\tilde{x})}{\partial\tilde{x}} + (1-\beta)\frac{\partial\eta(\tilde{x})}{\partial\tilde{x}}(\Delta(1;\tilde{x}) - \Delta(0;\tilde{x})) + (1-\beta)\eta(\tilde{x})\frac{\partial\Delta(1;\tilde{x})}{\partial\tilde{x}} + (1-\eta(\tilde{x})\beta\eta(\tilde{x}))\frac{\partial\Delta(0;\tilde{x})}{\partial\tilde{x}}.$$

Since $g(\cdot)$ has the monotone likelihood ratio property, $\eta(\tilde{x})$ is increasing in $\tilde{x}$. Now observe that $\mu^*(1,0;\tilde{x})$ is increasing in $\tilde{x}$ and $\mu^*(0,0;\tilde{x})$ is decreasing in $\tilde{x}$, which means that $\Delta(0)$ is increasing in $\tilde{x}$. Moreover, $\mu^*(1,1;\tilde{x})$ is increasing in $\tilde{x}$ so $\Delta(1)$ is increasing in $\tilde{x}$. Further, by definition of posterior beliefs we have $\Delta(1;\tilde{x}) \geq \Delta(0;\tilde{x})$ so this expression is increasing in $\tilde{x}$. Hence by the intermediate value theorem there is a unique $\tilde{x}^*$ solving Equation 1 such that the incompetent politician plays $a = 1$ when $x^0 > \tilde{x}^*$ and plays $a = 0$ when $x^0 \leq \tilde{x}^*$. ■

Corollary A.1 *The equilibrium cutoff $\tilde{x}^*$ is increasing in β.*

Proof of Corollary A.1: Using the definition of the cutoff $\tilde{x}^*$, define the function $I(\tilde{x})$ as

$$I(\tilde{x}) := 2\eta(\tilde{x}) - 1 + (1-\beta)\eta(\tilde{x})\Delta(1;\tilde{x}) + (1-\eta(\tilde{x}) + \beta\eta(\tilde{x}))\Delta(0;\tilde{x}),$$

and note that the equilibrium cutoff is defined by the value $\tilde{x}^*$ such that $I(\tilde{x}^*) = 0$. Further observe that, by definition of the existence of the equilibrium cutoff, $\frac{\partial I(\tilde{x})}{\partial \tilde{x}} > 0$. By the implicit function theorem,

$$\frac{d\tilde{x}^*}{d\beta} = -\frac{\partial I(\tilde{x})/\partial \beta}{\partial I(\tilde{x})\ \partial \tilde{x}}.$$

Partially differentiating with respect to β yields

$$\frac{\partial I(\tilde{x})}{\partial \beta} = -\eta(\tilde{x})\Big(\Delta(1;\tilde{x}) - \Delta(0;\tilde{x})\Big) + (1 - \eta(\tilde{x}) + \beta\eta(\tilde{x}))\frac{\partial \Delta(0;\tilde{x})}{\partial \beta}.$$

Now observe that

$$\frac{\partial \mu^*(1,0;\tilde{x})}{\partial \beta} = -\frac{\tau\pi(1-\tau)(1-\pi)(1-G(\tilde{x};0))}{(\tau\pi + \beta(1-\pi)(1-\tau)(1-G(\tilde{x};0)) + \pi(1-\tau)(1-G(\tilde{x};1)))^2} < 0,$$

and

$$\frac{\partial \mu^*(0,0;\tilde{x})}{\partial \beta} = \frac{\tau\pi(1-\pi)(1-\tau)G(\tilde{x};1)}{(\beta(1-\pi)(\tau + (1-\tau)G(\tilde{x};0) + \pi(1-\tau)G(\tilde{x};1))^2} > 0.$$

Therefore $\frac{\partial \Delta(0)}{\partial \beta} < 0$. Hence $\frac{\partial I(\tilde{x})}{\partial \beta} < 0$ so by the implicit function theorem, $\frac{d\tilde{x}^*}{d\beta} > 0$. ■

Lemma A.2 *Given an equilibrium cutoff $\tilde{x}^*$, there exists an optimal $\beta^* \in [0,1]$.*

Proof of Lemma A.2: The special interest group's objective function is

$$\max_{\beta\in[0,1]}\ 1 - \tau\pi - (1-\tau)\pi(1 - G(\tilde{x}^*(\beta);1)) - (1-\tau)(1-\pi)(1-G(\tilde{x}^*(\beta);0)) - c(\beta).$$

Differentiating with respect to β yields the first-order condition

$$(1-\tau)\pi g(\tilde{x}^*(\beta);1)\frac{d\tilde{x}^*}{d\beta} + (1-\tau)(1-\pi)g(\tilde{x}^*(\beta);0)\frac{d\tilde{x}^*}{d\beta} - c'(\beta) = 0.$$

Since the objective function is continuous and β is maximized along a compact interval, it must have both a maximum and a minimum. It is clear that the first-order condition must have at least one solution, as rearranging gives

$$(1-\tau)\frac{d\tilde{x}^*}{d\beta}\Big(\pi g(\tilde{x}^*;1) + (1-\pi)g(\tilde{x}^*;0)\Big) = c'(\beta),$$

but this solution may characterize either a maximum or a minimum. A maximum is characterized whenever the second-order condition is negative at the solution to the above first-order condition. The second order condition is

$$SOC = (1-\tau)\pi g(\tilde{x}^*;1)\frac{d^2\tilde{x}^*}{d\beta^2} + (1-\tau)\pi g'(\tilde{x}^*;1)(\frac{d\tilde{x}^*}{d\beta})^2 + (1-\tau)(1-\pi)g(\tilde{x}^*;0)\frac{d^2\tilde{x}^*}{d\beta^2}$$

$$+ (1-\tau)(1-\pi)g'(\tilde{x}^*;0)(\frac{d\tilde{x}^*}{d\beta})^2 - c''(\beta).$$

$$= (1-\tau)(\frac{d\tilde{x}^*}{d\beta})^2(\pi g'(\tilde{x}^*;1) + (1-\pi)g'(\tilde{x}^*;0)) + \frac{d^2\tilde{x}^*}{d\beta^2}(\frac{d\tilde{x}^*}{d\beta})^{-1}c'(\beta) - c''(\beta),$$

where the simplification follows from substituting from the first-order condition that $\pi g(\tilde{x}^*;1) + (1-\pi)g(\tilde{x}^*;0) = \frac{c'(\beta)}{1-\tau}(\frac{d\tilde{x}^*}{d\beta})^{-1}$.

The second-order condition is not readily globally concave: the sign of the first term depends on the value of $\tilde{x}^*$ by log-concavity of $g(\cdot)$, the second term depends on the sign of $\frac{d^2\tilde{x}^*}{d\beta^2}$, and the third term is negative. But note that if the second-order condition fails at the critical point, the maximum of the objective function must be on the corner. Further, from the first-order condition, observe that if $c'(\beta) \to 0$, the LHS is strictly positive and so the optimal solution is a corner solution at $\beta^* = 1$. If $c'(\beta)$ is relatively large, the LHS is strictly negative and the optimal solution is a corner solution at $\beta^* = 0$. ■

Proof of Proposition 1: Immediate from Lemmas A.1 and A.2. ■

Domestic Politics: Results

This section proves Results 1 and 2.

Recall from the main text that the probability of climate action is written as

$$A(\tilde{x}^*) = \tau\pi + (1-\tau)\pi(1 - G(\tilde{x}^*;1)) + (1-\tau)(1-\pi)(1 - G(\tilde{x}^*;0)).$$

Proof of Result 1: Differentiating with respect to β yields

$$\frac{dA(\tilde{x}^*)}{d\beta} = -(1-\tau)\pi g(\tilde{x}^*;1)\frac{d\tilde{x}^*}{d\beta} - (1-\tau)(1-\pi)g(\tilde{x}^*;0)\frac{d\tilde{x}^*}{d\beta} < 0.$$

■

Before proving Result 2, I prove a result about the cutoff $\tilde{x}^*$.

Lemma A.3 *The following are true about the incompetent politician's equilibrium cutoff:*

1. $\lim_{\pi\to 0} \tilde{x}^* = \infty$.
2. $\lim_{\pi\to 1} \tilde{x}^* = -\infty$.

Proof of Lemma A.3:

1. It is immediate that when $\pi \to 0$, we have $\eta(x) \to 0$ for any x. Then

$$\lim_{\pi\to 0} I(\tilde{x}) = -1 + \lim_{\pi\to 0} \Delta(0;\tilde{x}),$$

where the second term is less than 1 given the definition of the posterior beliefs induced by any $\tilde{x}$. Hence $\lim_{\pi\to 0} I(\tilde{x}) < 0$, which means it is never optimal for the incompetent politician to choose $a = 1$, meaning $\tilde{x}^* \to \infty$.

2. It is immediate that when $\pi \to 1$, we have $\eta(x) \to 1$ for any x. Then

$$\lim_{\pi\to 1} I(\tilde{x}) = 1 + (1-\beta)\Delta(1;\tilde{x}) + \beta\Delta(0;\tilde{x}),$$

where the first two terms are positive and the third term is at most -1 given the definition of the posterior beliefs induced by any $\tilde{x}$. Hence $\lim_{\pi\to 1} I(\tilde{x}) > 0$, which means it is always optimal for the incompetent politician to choose $a = 1$, meaning $\tilde{x}^* \to -\infty$.

■

Proof of Result 2: Define the function $I_\beta(\beta)$ as

$$I_\beta(\beta) := (1-\tau)\pi g(\tilde{x}^*(\beta);1)\frac{d\tilde{x}^*}{d\beta} + (1-\tau)(1-\pi)g(\tilde{x}^*(\beta);0)\frac{d\tilde{x}^*}{d\beta} - c'(\beta) = 0.$$

Observe that, by Lemma A.3, at $\pi = 0$ and $\pi = 1$, $I_\beta < 0$ for any $\beta > 0$ and $I_\beta = 0$ for $\beta = 0$ so it is optimal for the special interest group to be truthful, $\beta^* = 0$. Further by Rolle's theorem there must be a $\hat{\pi} \in (0,1)$ where $\frac{\partial I_\beta(\beta)}{\partial \pi} = 0$, meaning that β^* is nonmonotonic in π.

Partially differentiating yields

$$\begin{aligned}
\frac{\partial I_\beta(\beta)}{\partial \pi} &= (1-\tau)\Big[g(\tilde{x}^*;1)\frac{d\tilde{x}^*}{d\beta} + \pi g'(\tilde{x}^*;1)\frac{d\tilde{x}^*}{d\pi}\frac{d\tilde{x}^*}{d\beta} + \pi g(\tilde{x}^*;1)\frac{d^2\tilde{x}^*}{d\beta d\pi} \\
&\quad - g(\tilde{x}^*;0)\frac{d\tilde{x}^*}{d\beta} + (1-\pi)g'(\tilde{x}^*;0)\frac{d\tilde{x}^*}{d\pi}\frac{d\tilde{x}^*}{d\beta} + (1-\pi)g(\tilde{x}^*;0)\frac{d^2\tilde{x}^*}{d\beta d\pi}\Big]. \\
\Leftrightarrow \frac{\partial I_\beta(\beta)}{\partial \pi} &= (1-\tau)\Bigg[\Big(g(\tilde{x}^*;1) - g(\tilde{x}^*;0)\Big)\frac{d\tilde{x}^*}{d\beta} + \Big(\pi g'(\tilde{x}^*;1) + (1-\pi)g'(\tilde{x}^*;0)\Big)\frac{d\tilde{x}^*}{d\pi}\frac{d\tilde{x}^*}{d\beta} \\
&\quad + \Big(\pi g(\tilde{x}^*;1) + (1-\pi)g(\tilde{x}^*;0)\Big)\frac{d^2\tilde{x}^*}{d\beta d\pi}\Bigg].
\end{aligned}$$

Observe that at $\pi = 0$ and $\pi = 1$, $\frac{\partial I_\beta(\beta)}{\partial \pi} = 0$, implying that such points are extrema, and we know that $\beta^* = 0$ in these cases. But because $\beta \in [0,1]$, these must be minima. Then the point $\hat{\pi}$ which is defined by Rolle's theorem must be an interior maximum such that β^* is increasing when $\pi < \hat{\pi}$ and decreasing when $\pi > \hat{\pi}$. Such a $\hat{\pi}$ is characterized by $\frac{\partial I_\beta(\beta)}{d\pi} = 0$ and $\frac{\partial^2 I_\beta(\beta)}{d\pi^2} \leq 0$. ■

Corollary A.2 *The equilibrium signal cutoff $\tilde{x}^*$ is decreasing in π.*

Proof of Corollary A.2: By the implicit function theorem,

$$\frac{d\tilde{x}^*}{d\pi} = -\frac{\partial I(\tilde{x})/\partial \pi}{\partial I(\tilde{x})\ \partial \tilde{x}}.$$

Partially differentiating with respect to π yields

$$\frac{\partial I(\tilde{x})}{\partial \pi} = 2\frac{\partial \eta(\tilde{x})}{\partial \pi} + (1-\beta)\frac{\partial \eta(\tilde{x})}{\partial \pi}\Big(\Delta(1;\tilde{x}) - \Delta(0;\tilde{x})\Big) + (1-\eta(\tilde{x}) + \beta\eta(\tilde{x}))\frac{\partial \Delta(0;\tilde{x})}{\partial \pi}.$$

Now, $\frac{\partial \eta(\tilde{x})}{\partial \pi} = \frac{g(\tilde{x};1)g(\tilde{x};0)}{((1-\pi)G(\tilde{x};0)-\pi G(\tilde{x};1))^2} > 0$, $\frac{\partial \mu^*(1,0;\tilde{x})}{\partial \pi} = \frac{\beta\tau(1-\tau)(1-G(\tilde{x};0))}{(-\pi+\beta(-1+\pi+\tau-\tau\pi)+(1-\tau)(\beta(1-\pi)G(\tilde{x};0)+\pi G(\tilde{x};1)))^2} > 0$ and $\frac{\partial \mu^*(0,0;\tilde{x})}{\partial \pi} = -\frac{\tau\beta(1-\tau)G(\tilde{x};1)}{(\beta(1-\pi)(\tau+(1-\tau)G(\tilde{x};0))+\pi(1-\tau)G(\tilde{x};1))^2} < 0$. Then $\frac{\partial \Delta(0)}{\partial \pi} > 0$ and $\frac{\partial I(\tilde{x})}{\partial \pi} > 0$ so by the implicit function theorem, $\frac{d\tilde{x}^*}{d\pi} < 0$. ■

Extension: Pro-Climate Interest Group

Suppose that instead of a special interest group biased against climate action, the group that disseminates information to the voter is in favor of ambitious climate policies. Specifically, the group designs a signal $s \in \{0,1\}$ according to the experiment

$$\mathcal{E}(s=0,\omega=0) = 1-\gamma. \ \ \mathcal{E}(s=1,\omega=0) = \gamma.$$
$$\mathcal{E}(s=0,\omega=1) = 0. \ \ \mathcal{E}(s=1,\omega=1) = 1.$$

The interest group therefore chooses the parameter $\gamma \in [0,1]$. All of the analysis remains as before. I characterize the equilibrium of the climate policy subgame and show that the special interest's optimal choice of γ exists, as in the main text for an anti-climate interest group.

Lemma A.4 *A unique cutoff $\tilde{x}^*$ exists, admitting a unique perfect Bayesian equilibrium to the climate policy subgame with a pro-climate interest group. A politician of type θ chooses policy $a=1$ given signal x^θ with probability $\sigma^*(\theta,x^\theta) \in [0,1]$. These probabilities are*

$$\sigma^*(1,x^1) = x^1 = \omega.$$
$$\sigma^*(0,x^0) = 1 - G(\tilde{x}^*;\omega).$$

Upon observing policy a and signal s, the voter reelects the politician with probability $F(\mu^*(a,s;\tilde{x}^*))$.

Proof of Lemma A.4: The voter observes (a,s) and retains the politician when $\mu(a,s) \geq \varepsilon$, occurring with probability $F(\mu(a,s))$.

The competent politician always follows her signal. If $x^1 = 1$, playing $a=1$ is optimal:

$$1 + F(\mu(1,1)) \geq F(\mu(0,1)) \ \Leftrightarrow \ 1 \geq -\Delta(1).$$

Similarly, if $x^1 = 0$, the competent politician chooses $a = 0$:

$$\gamma F(\mu(1,1)) + (1-\gamma)F(\mu(1,0)) \leq 1 + \gamma F(\mu(0,1)) + (1-\gamma)F(\mu(0,0)) \Leftrightarrow \gamma\Delta(1) + (1-\gamma)\Delta(0) \leq 1.$$

Given the signal $x^0 = x$, the incompetent type chooses $a = 1$ iff

$$\begin{aligned}
\eta(x) + \gamma(1-\eta(x))F(\mu(1,1)) + (1-\gamma)(1-\eta(x))F(\mu(1,0)) + \eta(x)F(\mu(1,1) \geq \\
(1-\eta(x)) + \gamma(1-\eta(x))F(\mu(0,1)) + (1-\gamma)(1-\eta(x))F(\mu(0,0)) + \eta(x)F(\mu(0,1)) \\
\Leftrightarrow\ 2\eta(x) - 1 + (\gamma - \gamma\eta(x) + \eta(x))\Delta(1) + (1-\gamma)(1-\eta(x))\Delta(0) \geq 0.
\end{aligned}$$

Let $\tilde{x}$ be the value of x that solves this at equality. The posterior beliefs induced by these strategies are

$$\begin{aligned}
\mu(1,0) &= 0. \\
\mu(1,1) &= \frac{\tau\pi\frac{\gamma}{1-\gamma}}{\pi\frac{\gamma}{1-\gamma}(\tau + (1-\tau)(1-G(\tilde{x};1))) + (1-\pi)(1-G(\tilde{x};0))}. \\
\mu(0,0) &= \frac{\tau}{\tau + (1-\tau)G(\tilde{x};0)}. \\
\mu(0,1) &= \frac{\tau(1-\pi)}{(1-\pi)(\tau + (1-\tau)G(\tilde{x};0)) + (1-\tau)\pi\frac{\gamma}{1-\gamma}G(\tilde{x};1)}.
\end{aligned}$$

Differentiating the incompetent type's constraint with respect to x yields

$$2 - \frac{\partial\eta(x)}{\partial x}(1-\gamma)(\Delta(1) + \Delta(0)) + (\gamma(1-\eta(x)) + \eta(x))\frac{\partial\Delta(1)}{\partial x} + (1-\eta(x))(1-\gamma)\frac{\partial\Delta(0)}{\partial x}.$$

Since $\mu(0,0)$ is decreasing in $\tilde{x}$, $\Delta(0)$ is increasing in $\tilde{x}$. Similarly, $\mu(1,1)$ is increasing in $\tilde{x}$ and $\mu(0,1)$ is decreasing in $\tilde{x}$. Hence by the intermediate value theorem there is a $\tilde{x}^*$ such that the incompetent type plays $a = 1$ iff $x^0 \geq \tilde{x}^*$. ■

Corollary A.3 *The equilibrium signal cutoff $\tilde{x}$ is decreasing in γ in the model with a pro-climate interest group.*

Proof of Corollary A.3: Define the function

$$I_\gamma(\tilde{x}) := 2\eta(\tilde{x}) - 1 + (\gamma - \gamma\eta(\tilde{x}) + \eta(\tilde{x}))\Delta(1;\tilde{x}) + (1-\gamma)(1-\eta(\tilde{x}))\Delta(0;\tilde{x}).$$

Clearly, $I_\gamma(\tilde{x})$ is increasing in $\tilde{x}$ and the point $\tilde{x}^*$ is defined by $I_\gamma(\tilde{x}^*) = 0$. By the implicit function theorem,

$$\frac{d\tilde{x}^*}{d\gamma} = -\frac{\partial I_\gamma(\tilde{x})/\partial\gamma}{\partial I_\gamma(\tilde{x})/\partial\tilde{x}}.$$

Differentiating with respect to γ yields

$$\frac{\partial I_\gamma(\tilde{x})}{\partial \gamma} = (1-\eta(\tilde{x}))(\Delta(1;\tilde{x}) - \Delta(0;\tilde{x})) + (\gamma - \gamma\eta(\tilde{x}) + \eta(\tilde{x}))\frac{\partial \Delta(1;\tilde{x})}{\partial \gamma}.$$

Now, $\mu(1,1)$ is increasing in γ and $\mu(1,0)$ is decreasing in γ so $\frac{\partial \Delta(1;\tilde{x})}{\partial \gamma} > 0$. Hence $\frac{\partial I_\gamma(\tilde{x})}{\partial \gamma} > 0$ so by the implicit function theorem $\frac{d\tilde{x}^*}{d\gamma} < 0$. ■

Lemma A.5 *Given equilibrium behavior in the climate policy subgame with a pro-climate interest group, there exists an optimal* $\gamma^* \in [0,1]$.

Proof of Lemma A.5: The special interest then chooses γ to maximize the probability of climate action, given by the objective function

$$\max_{\gamma \in [0,1]} \tau\pi + (1-\tau)\pi(1-G(\tilde{x}^*(\gamma);1)) + (1-\tau)(1-\pi)(1-G(\tilde{x}^*(\gamma);0)) - c(\gamma).$$

Differentiating with respect to γ yields the (rearranged) first-order condition

$$-(1-\tau)\pi g(\tilde{x}^*;1)\frac{d\tilde{x}^*}{d\gamma} - (1-\tau)(1-\pi)g(\tilde{x}^*;0)\frac{d\tilde{x}^*}{d\gamma} = c'(\gamma)$$

and second-order condition

$$-(1-\tau)\frac{d^2\tilde{x}^*}{d\gamma^2}\Big(\pi g(\tilde{x}^*;1) + (1-\pi)g(\tilde{x}^*;0)\Big) - (1-\tau)(\frac{d\tilde{x}^*}{d\gamma})^2\Big(\pi g'(\tilde{x}^*;1) + (1-\pi)g'(\tilde{x}^*;0)\Big) - c''(\gamma),$$

with characterization analogous to the proof in Lemma A.2. ■

Extension: Politician and Interest Group Signal

In the main model, the politician is unable to condition her strategy on the signal s sent by the interest group. I now relax that assumption. This means that the politician's strategy is now a function of her type θ, her private signal x^θ, as well as the public signal s.

It is straightforward to observe that the interest group's signal has no effect on the competent type: since she knows ω perfectly already, there is no incentive to deviate from her equilibrium strategy as posited in the main text. Hence, $\sigma^*(1, x^1, s) = x^1 = \omega$.

Now consider the incompetent type. Define $\rho(x^0, s; \beta) = P(\omega = 1 | x^0, s; \beta)$ to be the incompetent type's posterior belief that $\omega = 1$ given her private signal x^0 and the realization of the interest group's message s given β. Since $s = 1$ is a truthful message, $\rho(x^0, 1) = 1$ for any value of x^0. Hence in the subgame following $s = 1$, the incompetent type chooses $a = 1$ iff

$$1 + F(\mu(1,1)) \geq F(\mu(0,1)),$$

which is always satisfied, so $\sigma^*(0, x^0, 1) = 1$. Hence following $s = 1$, there is a pooling equilibrium on $a = 1$. Off path, $\mu(0,1) = 0$ as it is the incompetent type who would possibly deviate.

Following $s = 0$, the incompetent type does not know if the special interest is truthful reporting $\omega = 0$ or if with some probability β it misreported. Her posterior belief is $\rho(x^0, 0; \beta) = \frac{g(x^0;1)\beta\pi}{g(x^0;1)\beta\pi + g(x^0;0)(1-\pi)}$. Then the incompetent type's problem is to choose $a = 1$ whenever

$$\rho(x^0, 0; \beta) + F(\mu(1, 0)) \geq (1 - \rho(x^0, 0; \beta)) + F(\mu(0, 0)) \;\Leftrightarrow\; 2\rho(x^0, 0; \beta) - 1 + \Delta(0) \geq 0.$$

It is clear that $\mu(1, 0) = 0$ but $\mu(0, 0)$ is decreasing in the probability that the incompetent type takes action (analogous to proof of Lemma A.1). Then there is a cutoff $\tilde{x}^*$ such that the incompetent type plays $a = 1$ iff $x^0 \geq \tilde{x}^*$. Hence $\sigma^*(0, x^0, 0) = 1 - G(\tilde{x}^*; \omega)$.

Now observe that $\frac{\partial \rho(x^0,0;\beta)}{\partial \beta} = \frac{\pi(1-\pi)g(x^0;1)g(x^0;0)}{(g(x^0;1)\pi\beta + g(x^0;0)(1-\pi))^2} > 0$ for any β and any signal x^0, as the politician rationally discounts the special interest group's signal as reporting is more likely to be biased. Moreover, $\frac{\partial \mu(0,0)}{\partial \beta} > 0$ from Corollary A.1. The result that $\frac{d\tilde{x}^*}{d\beta} > 0$ holds if $\frac{\partial \rho(x^0,0;\beta)}{\partial \beta} < \frac{\partial \mu(0,0)}{\partial \beta}$, or the voter's posterior belief about the politician's competence is more responsive to special interest reporting than the incompetent politician's posterior belief about the state of the world. Such a condition is both theoretically plausible within the context of the model as the politician has more information at her disposal than the voter does. Additionally, the incompetent politician, in seeking to be reelected, would prefer that the voter has a higher assessment of her competence.

Proofs of International Cooperation Model

International Cooperation: Equilibrium

I prove Proposition 2 in a series of lemmas. To prove equilibrium existence in the subgame, I proceed in several steps. First I assume that competent types are willing to follow their signal, $a_i = x_i^1$ to prove that incompetent types have equilibrium strategies that are cutoffs: incompetent types play a_i iff their signal x_i^0 exceeds some cutoff k_i, holding fixed the behavior in country j and the beliefs of voter i (Lemma A.7). I then characterize the conditions needed for competent types to follow their signal in equilibrium, namely when τ_i is greater than some threshold $\overline{\tau}_i$ (Lemma A.8). I make this assumption for the rest of the analysis (Assumption A.1). I then endogenize the behavior of politician j into incompetent politician i's cutoff to show that there are cutoffs that are mutual best responses (Lemma A.9). Finally, I endogenize voter i's posterior beliefs as a function of the cutoff strategies (Lemma A.10). I then state existence of the optimal misreporting levels in Lemmas A.12 and A.13.

Recall that $\hat{\sigma}_i(\theta_i, \omega)$ is voter i's belief about the probability that politician i chooses $a_i = 1$ when she is of type θ and the state of the world is ω. Define $B(\theta_i, a_i, s_i, a_j)$ as proportional to the *ex ante* probability that a politician i of type θ chooses action a_i and signal s_i is realized by the special interest group in country i and politician j chooses action a_j.

$$B(\theta_i, 1, s_i, a_j) = P(\theta_i)\Big(\pi\hat{\sigma}_i(\theta_i, 1)P(a_j|\omega = 1) + (1 - \pi)\hat{\sigma}_i(\theta_i, 0)\frac{P(s_i|s_i \neq \omega)}{P(s_i|s_i = \omega)}P(a_j|\omega = 0)\Big).$$

$$B(\theta_i, 0, s_i, a_j) = P(\theta_i)\Big(\pi(1-\hat{\sigma}_i(\theta_i, 1))P(a_j|\omega=1) + (1-\pi)(1-\hat{\sigma}_i(\theta_i, 0))\frac{P(s_i|s_i \neq \omega)}{P(s_i|s_i = \omega)}P(a_j|\omega=0)\Big).$$

Upon observing politician i's policy a_i, the special interest's signal in country i s_i, and politician j's policy a_j, voter i has a posterior belief about politician i's competence $\mu_i(a_i, s_i, a_j) = P(\theta_i = 1|a_i, s_i, a_j)$,

$$\mu_i(a_i, s_i, a_j) = \frac{P(a_i, s_i, a_j|\theta_i = 1)P(\theta_i = 1)}{P(a_i, s_i, a_j|\theta_i = 1)P(\theta_i = 1) + P(a_i, s_i, a_j|\theta_i = 0)P(\theta_i = 0)} = \frac{B(1, a_i, s_i, a_j)}{B(1, a_i, s_i, a_j) + B(0, a_i, s_i, a_j)}$$

Lemma A.6 *The following statements are true regarding the ordering of voter i's posterior beliefs:*

- $\mu_i(1, 1, a_j) \geq \mu_i(0, 1, a_j) \;\Leftrightarrow\; \hat{\sigma}_i(1, 1) \geq \hat{\sigma}_i(0, 1)$.
- $\mu_i(1, 0, 1) \geq \mu_i(0, 0, 1) \;\Leftrightarrow\; \frac{\beta_i}{1-\beta_i}\frac{\pi}{1-\pi} \geq \frac{\tau_j\hat{\sigma}_j(1,1)+(1-\tau_j)\hat{\sigma}_j(0,1)}{\tau_j\hat{\sigma}_j(1,0)+(1-\tau_j)\hat{\sigma}_j(0,0)}\frac{\hat{\sigma}_i(0,1)-\hat{\sigma}_i(1,1)}{\hat{\sigma}_i(1,0)-\hat{\sigma}_i(0,0)}$.
- $\mu_i(1, 0, 0) \geq \mu_i(0, 0, 0) \;\Leftrightarrow\; \frac{\beta_i}{1-\beta_i}\frac{\pi}{1-\pi} \geq \frac{\tau_j(1-\hat{\sigma}_j(1,1))+(1-\tau_j)(1-\hat{\sigma}_j(0,1))}{\tau_j(1-\hat{\sigma}_j(1,0))+(1-\tau_j)(1-\hat{\sigma}_j(0,0))}\frac{\hat{\sigma}_i(0,1)-\hat{\sigma}_i(1,1)}{\hat{\sigma}_i(1,0)-\hat{\sigma}_i(0,0)}$.
- $\mu_i(1, 0, 1) \geq \mu_i(1, 0, 0) \;\Leftrightarrow\; \frac{\hat{\sigma}_i(1,1)}{\hat{\sigma}_i(0,1)} \geq \frac{\hat{\sigma}_i(1,0)}{\hat{\sigma}_i(0,0)}$.
- $\mu_i(1, 1, a_j) \geq \mu_i(1, 0, a_j) \;\Leftrightarrow\; \frac{\hat{\sigma}_i(1,1)}{\hat{\sigma}_i(0,1)} \geq \frac{\hat{\sigma}_i(1,0)}{\hat{\sigma}_i(0,0)}$.
- $\mu_i(0, 1, a_j) \geq \mu_i(0, 0, a_j) \;\Leftrightarrow\; \frac{1-\hat{\sigma}_i(1,1)}{1-\hat{\sigma}_i(0,1)} \geq \frac{1-\hat{\sigma}_i(1,0)}{1-\hat{\sigma}_i(0,0)}$.
- $\mu_i(0, 1, 1) \geq \mu_i(0, 1, 0) \;\Leftrightarrow\; \frac{\hat{\sigma}_i(1,1)}{\hat{\sigma}_i(0,1)} \geq \frac{\hat{\sigma}_i(1,0)}{\hat{\sigma}_i(0,0)}$.
- $\mu_i(0, 0, 1) \geq \mu_i(0, 0, 0) \;\Leftrightarrow\; \frac{1-\hat{\sigma}_i(1,1)}{1-\hat{\sigma}_i(0,1)} \geq \frac{1-\hat{\sigma}_i(1,0)}{1-\hat{\sigma}_i(0,0)}$.
- $\mu_i(a_i, 1, 1) = \mu_i(a_i, 1, 0)$.

Proof of Lemma A.6: Straightforward from definition of posterior beliefs. ■

Write $\Delta_i(s_i, a_j) = F(\mu_i(1, s_i, a_j)) - F(\mu_i(0, s_i, a_j))$ as the difference in politician i's re-election probabilities from playing $a_i = 1$ and $a_i = 0$ when interest group i generates signal s_i and politician j plays a_j.

Corollary A.4 *Suppose competent politicians follow their signal,* $\hat{\sigma}_i(1, 1) = 1$ *and* $\hat{\sigma}_i(1, 0) = 0$. *The following statements are true:*

- $\Delta_i(1, a_j) \geq 0$.
- $\Delta_i(1, a_j) \geq \Delta_i(0, a_j)$.
- $\Delta_i(s_i, 1) \geq \Delta_i(s_i, 0)$.

Proof of Corollary A.4: Immediate from Lemma A.6 and the definition of $\Delta_i(s_i, a_j)$. ■

Define the set of domestic fundamentals, which is the set of country-specific electoral returns in the climate policy subgame, as $\Lambda_i = \Big(\Delta_i(1,1),\ \Delta_i(1,0),\ \Delta_i(0,1),\ \Delta_i(0,0)\Big)$.

Lemma A.7 *Fix Λ_i and assume that competent politicians follow their signal. Politician i's best response is in cutoff strategies: there exists a cutoff k_i such that the incompetent politician chooses $a_i = 1$ iff $x_i^0 \geq k_i$.*

Proof of Lemma A.7: Let $y_j = P(a_j = 1|x_i)$ be the probability that politician j chooses $a_j = 1$ given what politician i knows about the state of the world from her realized signal $x_i^0 = x_i$. Then, $y_j = \tau_j \eta(x_i) + (1-\tau_j)\eta(x_i)\hat{\sigma}_j(0,1) + (1-\tau_j)(1-\eta(x_i))\hat{\sigma}_j(0,0)$. Observe that y_j is increasing in $\eta(x_i)$ (and hence in x_i by the monotone likelihood ratio property): $\frac{\partial y_j}{\partial \eta(x_i)} = \tau_j + (1-\tau_j)(\hat{\sigma}_j(0,1) - \hat{\sigma}_j(0,0)) \geq 0$, which follows from monotonicity of strategies in the state of the world.

Given Λ_i, the incompetent politician plays $a_i = 1$ following signal x_i iff

$$\begin{aligned}
&\eta(x_i)\Bigg[y_j\Big[1 + \beta_i F(\mu_i(1,0,1)) + (1-\beta_i)F(\mu_i(1,1,1)\Big] + (1-y_j)\Big[\beta_i F(\mu_i(1,0,0)) + (1-\beta_i)F(\mu_i(1,1,0))\Big]\Bigg] \\
&\qquad + (1-\eta(x_i))\Bigg[y_j\Big[1 + F(\mu_i(1,0,1))\Big] + (1-y_j)F(\mu_i(1,0,0))\Bigg] \geq \\
&\eta(x_i)\Bigg[y_j\Big[\beta_i F(\mu_i(0,0,1)) + (1-\beta_i)F(\mu_i(0,1,1))\Big] + (1-y_j)\Big[1 + \beta_i F(\mu_i(0,0,0)) + (1-\beta_i)F(\mu_i(0,1,0))\Big]\Bigg] \\
&\qquad + (1-\eta(x_i))\Bigg[y_j F(\mu_i(0,0,1)) + (1-y_j)\Big[1 + F(\mu_i(0,0,0))\Big]\Bigg]. \\
&\Leftrightarrow\ 2y_j - 1 + (1-\eta(x_i) + \eta(x_i)\beta_i)\Big(y_j\Delta_i(0,1) + (1-y_j)\Delta_i(0,0)\Big) + \eta(x_i)(1-\beta_i)\Big(y_j\Delta_i(1,1) + (1-y_j)\Delta_i(1,0)
\end{aligned} \tag{2}$$

Differentiating with respect to $\eta(x_i)$ yields

$$2\frac{\partial y_j}{\partial \eta(x_i)} + (1-\beta_i)\Big(y_j\Delta_i(1,1) - y_j\Delta_i(0,1) + (1-y_j)\Delta_i(1,0) - (1-y_j)\Delta_i(0,0)\Big) + (1-\eta(x_i)+\eta(x_i)\beta_i)\frac{\partial y_j}{\partial \eta(x_i)}\Big(\Delta_i(0,1) - \Delta_i(0,0)\Big) >$$

Hence, incompetent politician i's net gain from playing $a_i = 1$ is increasing in x_i such that by the intermediate value theorem she adopts a cutoff strategy and plays $a_i = 1$ iff $x_i^0 \geq k_i$. ■

Lemma A.8 *Assume $\tau_j \geq \overline{\tau}_j$. The competent politician i always follows her signal.*

Proof of Lemma A.8: Proof is analogous for politician j. Let $y_{j\omega} = P(a_j = 1|\omega)$ be the competent politician's updated beliefs about politician j's behavior given that she knows the state of the world perfectly.

Suppose a competent politician i observes $x_i = 1$. She plays $a_i = 1$ iff

$$y_{j1}\Big[1+\beta_i F(\mu_i(1,0,1)) + (1-\beta_i)F(\mu_i(1,1,1))\Big] + (1-y_{j1})\Big[\beta_i F(\mu_i(1,0,0)) + (1-\beta_i)F(\mu_i(1,1,0))\Big] \geq$$
$$y_{j1}\Big[\beta_i F(\mu_i(0,0,1)) + (1-\beta_i)F(\mu_i(0,1,1))] + (1-y_{j1})\Big[\beta_i F(\mu_i(0,0,0)) + (1-\beta_i)F(\mu_i(0,1,0))\Big].$$
$$\Leftrightarrow\ y_{j1} + \beta_i\Big(y_{j1}\Delta_i(0,1) + (1-y_{j1})\Delta_i(0,0)\Big) + (1-\beta_i)\Big(y_{j1}\Delta_i(1,1) + (1-y_{j1})\Delta_i(1,0)\Big) \geq 0.$$

By following her signal, $\Delta_i(1,1) = \Delta_i(1,0)$ and $\Delta_i(0,1) \geq \Delta_i(0,0)$ by Corollary A.4; the inequality holds.

Similarly, suppose a competent politician i observes $x_i = 0$. She plays $a_i = 0$ iff

$$y_{j0}F(\mu_i(0,0,1)) + (1-y_{j0})\Big[1 + F(\mu_i(0,0,0))\Big] \geq y_{j0}F(\mu_i(1,0,1)) + (1-y_{j0})F(\mu_i(1,0,0))$$
$$\Leftrightarrow\ (1-y_{j0})(1-\Delta_i(0,0)) - y_{j0}\Delta_i(0,1) \geq 0.$$

This inequality need not hold; by way of contradiction, suppose it doesn't hold. Then this means that the competent type plays $a_i = 1$ regardless of her signal. Therefore, $\mu_i(0, s_i, a_j) = 0$ because any $a_i = 0$ must be played by the incompetent type.

Consider the incentives for the incompetent type. Suppose the incompetent type receives signal $x_i^0 = x_i$ and has beliefs $y_j = P(a_j = 1|x_i)$. The incompetent type plays $a_i = 1$ iff Equation 2 is satisfied, which in this case reduces to

$$2y_j - 1 - F(0) + (\eta(x_i)\beta_i y_j + (1-\eta(x_i))y_j)F(\mu_i(1,0,1)) + \eta(x_i)(1-\beta_i)F(\mu_i(1,1,1)) + (\eta(x_i)\beta_i(1-y_j) + (1-\eta(x_i))(1-y_j))F(\mu_i(1,0,$$

where the simplification comes from the fact that $\mu_i(0, s_i, a_j) = 0$ and that by Lemma A.6, $\mu_i(1,1,1) = \mu_i(1,1,0)$. In a pooling equilibrium, it would also be true that $\mu_i(1, s_i, a_j) = \tau_i$, meaning the voter would learn nothing about the competent politician's type from $a_i = 1$. Substituting this into the incompetent politician's incentive constraint yields

$$2y_j - 1 - F(0) + F(\tau_i) \geq 0,$$

however, by the intermediate value theorem, there are some values of x_i where this constraint holds and some where it does not (because y_j is increasing in x_i). Hence pooling on $a_i = 1$ is not always optimal: there exists a cutoff $\hat{x}_i$ such that the incompetent politician plays $a_i = 1$ iff $x_i \geq \hat{x}_i$ and $a_i = 0$ otherwise. Then we know that in such an equilibrium, $\sigma(1,1) = \sigma(1,0) = 1$, $\sigma(0,1) = 1 - G(\hat{x}_i; 1)$, and $\sigma(0,0) = 1 - G(\hat{x}_i; 0)$. Further, by first-order stochastic dominance, $G(\hat{x}_i; 0) \geq G(\hat{x}_i; 1) \implies \sigma(0,1) \geq \sigma(0,0)$ so by Lemma A.6 we have $\mu_i(1,0,0;\hat{x}_i) \geq \mu_i(1,0,1;\hat{x}_i)$.

Returning the competent politician's constraint, recall that she plays $a_i = 1$ following $x_i^1 = 0$ iff

$$(1-y_{j0})(1 - F(\mu_i(1,0,0;\hat{x}_i))) - y_{j0}F(\mu_i(1,0,1;\hat{x}_i)) + F(0) \leq 0.$$

Note that if $x_i \geq \hat{x}_i$, then the incompetent politician is pooling, so $\mu_i(1,0,0;\hat{x}_i) = \mu_i(1,0,1;\hat{x}_i) = \tau_i$. But if $x_i \leq \hat{x}_i$, there is separation between the competent and the incompetent types, so posterior beliefs are bounded below by τ_i, $\mu_i(1,0,0;\hat{x}_i) \geq \mu_i(1,0,1;\hat{x}_i) \geq \tau_i$. Clearly this is hardest to satisfy at the lower bound, yielding

$$1 - y_{j0} - F(\tau_i) + F(0) \leq 0.$$

Since the LHS is increasing in τ_j and the RHS is constant, there is a value $\overline{\tau}_j$ such that when $\tau_j \leq \overline{\tau}_j$ the constraint is satisfied, but that contradicts the hypothesis that $\tau_j \geq \overline{\tau}_j$. ∎

To continue with the analysis, maintain the following assumption such that Lemma A.8 holds.

Assumption A.1 *Both politicians are sufficiently likely to be competent: $\tau_i \geq \overline{\tau}_i$ and $\tau_j \geq \overline{\tau}_j$.*

Now we endogenize the behavior of politician j into politician i's best response (and vice versa).

Lemma A.9 *Fix Λ_i. There exist cutoffs $(\tilde{x}_i, \tilde{x}_j)$ that are mutual best responses.*

Proof of Lemma A.9: By Lemma A.7, the incompetent politician in both countries has a well-defined cutoff k_i such that i plays $a_i = 1$ iff $x_i^0 \geq k_i$ holding fixed the strategy of politician j. Since politician j is playing a cutoff strategy, we know that $y_j = \tau_j\eta(x_i) + (1-\tau_j)\eta(x_i)(1-G(k_j;1)) + (1-\tau_j)(1-\eta(x_i))(1-G(k_j;0))$ given that $x_i^0 = x_i$. Observe that $\frac{\partial y_j}{\partial k_j} = -(1-\tau_j)\Big(\eta(x_i)g(k_j;1) + (1-\eta(x_i))g(k_j;0)\Big) < 0$.

Define $\hat{I}(k_i,k_j;\Lambda_i)$ as the incompetent politician i's indifference condition between playing $a_i = 1$ and $a_i = 0$ that endogenizes the cutoff of the incompetent politician j:

$$\begin{aligned}\hat{I}(k_i,k_j;\Lambda_i) := {}& 2y_j(k_i,k_j) - 1 + (1-\eta(k_i)+\eta(k_i)\beta_i)\Big(y_j(k_i,k_j)\Delta_i(0,1) + (1-y_j(k_i,k_j))\Delta_i(0,0)\Big) \\ &+ \eta(k_i)(1-\beta_i)\Big(y_j(k_i,k_j)\Delta_i(1,1) + (1-y_j(k_i,k_j))\Delta_i(1,0)\Big).\end{aligned}$$

Hence, the cutoffs $(\tilde{x}_i, \tilde{x}_j)$ therefore solve the system

$$\hat{I}(\tilde{x}_i,\tilde{x}_j;\Lambda_i) = 0 \quad \text{and} \quad \hat{I}(\tilde{x}_j,\tilde{x}_i;\Lambda_j) = 0.$$

To show the system has a unique solution, I use the implicit function theorem. The Jacobian matrix is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \hat{I}(k_i,k_j;\Lambda_i)}{\partial k_i} & \frac{\partial \hat{I}(k_i,k_j;\Lambda_i)}{\partial k_j} \\ \frac{\partial \hat{I}(k_j,k_i;\Lambda_j)}{\partial k_i} & \frac{\partial \hat{I}(k_j,k_i;\Lambda_j)}{\partial k_j} \end{bmatrix}.$$

Differentiating $\hat{I}(k_i,k_j;\Lambda_i)$ with respect to k_j yields

$$\frac{\partial \hat{I}(k_i,k_j;\Lambda_i)}{\partial k_j} = 2\frac{\partial y_j}{\partial k_j} + (1-\eta(k_i)+\eta(k_i)\beta_i)\frac{\partial y_j}{\partial k_j}\Big(\Delta_i(0,1) - \Delta_i(0,0)\Big) + \eta(k_i)(1-\beta_i)\frac{\partial y_j}{\partial k_j}\Big(\Delta_i(1,1) - \Delta_i(1,0)\Big) < 0.$$

Since the determinant $|\mathbf{J}| = \frac{\partial \hat{I}(k_i,k_j;\Lambda_i)}{\partial k_i}\frac{\partial \hat{I}(k_j,k_i;\Lambda_j)}{\partial k_j} - \frac{\partial \hat{I}(k_i,k_j;\Lambda_i)}{\partial k_j}\frac{\partial \hat{I}(k_j,k_i;\Lambda_j)}{\partial k_i} > 0$, the system has a unique solution at the cutoffs $(\tilde{x}_i, \tilde{x}_j)$ are well-defined. ■

Now we need to endogenize voter i's beliefs by writing $\Delta_i(s_i, a_j)$ as a function of equilibrium strategies. By Lemma A.9, there exists a pair of cutoffs $(\tilde{x}_i, \tilde{x}_j)$ such that incompetent politician i plays $a_i = 1$ iff $x_i^0 \geq \tilde{x}_i$ (same for incompetent politician j), and that competent politicians always follow their signals. Moreover, given $\tilde{x}_j$, we can write $\tilde{y}_{j1} = P(a_j = 1|\omega = 1, \tilde{x}_j) = \tau_j + (1-\tau_j)(1-G(\tilde{x}_j;1))$ and $\tilde{y}_{j0} = P(a_j = 1|\omega = 0, \tilde{x}_j) = (1-\tau_j)(1-G(\tilde{x}_j;0))$. This induces the following posterior beliefs for voter i:

$$\mu_i(1,0,1;\tilde{x}_i,\tilde{x}_j) = \frac{\tau_i \pi \tilde{y}_{j1}}{\pi\tilde{y}_{j1}(\tau_i + (1-\tau_i)(1-G(\tilde{x}_i;1))) + (1-\tau_i)(1-\pi)\frac{\beta_i}{1-\beta_i}(1-G(\tilde{x}_i;0))\tilde{y}_{j0}}.$$

$$\mu_i(1,1,1;\tilde{x}_i,\tilde{x}_j) = \frac{\tau_i}{\tau_i + (1-\tau_i)(1-G(\tilde{x}_i;1))}.$$

$$\mu_i(1,0,0;\tilde{x}_i,\tilde{x}_j) = \frac{\tau_i \pi (1-\tilde{y}_{j1})}{\pi(1-\tilde{y}_{j1})(\tau_i + (1-\tau_i)(1-G(\tilde{x}_i;1))) + (1-\tau_i)(1-\pi)\frac{\beta_i}{1-\beta_i}(1-G(\tilde{x}_i;0))(1-\tilde{y}_{j0})}.$$

$$\mu_i(1,1,0;\tilde{x}_i,\tilde{x}_j) = \frac{\tau_i}{\tau_i + (1-\tau_i)(1-G(\tilde{x}_i;1))}.$$

$$\mu_i(0,0,1;\tilde{x}_i,\tilde{x}_j) = \frac{\tau_i(1-\pi)\frac{\beta_i}{1-\beta_i}\tilde{y}_{j0}}{(1-\pi)\frac{\beta_i}{1-\beta_i}\tilde{y}_{j0}(\tau_i + (1-\tau_i)G(\tilde{x}_i;0)) + (1-\tau_i)\pi G(\tilde{x}_i;1)\tilde{y}_{j1}}.$$

$$\mu_i(0,1,1;\tilde{x}_i,\tilde{x}_j) = 0.$$

$$\mu_i(0,0,0;\tilde{x}_i,\tilde{x}_j) = \frac{\tau_i(1-\pi)\frac{\beta_i}{1-\beta_i}(1-\tilde{y}_{j0})}{(1-\pi)\frac{\beta_i}{1-\beta_i}(1-\tilde{y}_{j0})(\tau_i + (1-\tau_i)G(\tilde{x}_i;0)) + (1-\tau_i)\pi G(\tilde{x}_i;1)(1-\tilde{y}_{j1})}.$$

$$\mu_i(0,1,0;\tilde{x}_i,\tilde{x}_j) = 0.$$

The following table summarizes the sign of the derivative of voter i's posterior beliefs with respect to the cutoffs $\tilde{x}_i$ and $\tilde{x}_j$.

$\mu_i(a_i, s_i, a_j; \tilde{x}_i, \tilde{x}_j)$	$\frac{\partial \mu_i(a_i,s_i,a_j;\tilde{x}_i,\tilde{x}_j)}{\partial \tilde{x}_i}$	$\frac{\partial \mu_i(a_i,s_i,a_j;\tilde{x}_i,\tilde{x}_j)}{\partial \tilde{x}_j}$
$\mu_i(1,1,1)$	+	0
$\mu_i(1,1,0)$	+	0
$\mu_i(1,0,1)$	+	+
$\mu_i(1,0,0)$	+	-
$\mu_i(0,1,1)$	0	0
$\mu_i(0,1,0)$	0	0
$\mu_i(0,0,1)$	-	+
$\mu_i(0,0,0)$	-	-

Table A.1

Signing derivatives with respect to $\tilde{x}_i$ follow analogously from the proof of Lemma A.1.

Derivatives with respect to $\tilde{x}_j$ can be signed because the monotone likelihood ratio property implies hazard rate ordering. From Table A.1, it is evident that all $\Delta(s_i, a_j)$ are increasing in $\tilde{x}_i$.

Lemma A.10 *There exist cutoffs $(\tilde{x}_i^*, \tilde{x}_j^*)$ that are mutual best responses.*

Proof of Lemma A.10: Define $\hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)$ as the incompetent politician i's indifference condition that endogenizes both the cutoff of the incompetent politician j $\tilde{x}_j$ and the electoral returns in country i Λ_i given equilibrium strategies defined by such cutoffs:

$$\hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i) = 2y(\tilde{x}_i, \tilde{x}_j) - 1 + (1-\eta(\tilde{x}_i)+\eta(\tilde{x}_i)\beta_i)\Big(y(\tilde{x}_i, \tilde{x}_j)\Delta_i(0,1;\tilde{x}_i,\tilde{x}_j) + (1-y(\tilde{x}_i,\tilde{x}_j))\Delta_i(0,0;\tilde{x}_i,\tilde{x}_j)\Big) + \eta(\tilde{x}_i)(1-\beta_i)\Delta_i(1,0;\tilde{x}_i,\tilde{x}_j). \tag{3}$$

Differentiating with respect to $\tilde{x}_i$ yields

$$\frac{\partial \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)}{\partial \tilde{x}_i} = \frac{\partial \hat{I}(\tilde{x}_i; \tilde{x}_j, \Lambda_i)}{\partial \tilde{x}_i} + (1-\eta(\tilde{x}_i)-\eta(\tilde{x}_i)\beta_i)(y(\tilde{x}_i,\tilde{x}_j)\frac{\partial \Delta_i(0,1)}{\partial \tilde{x}_i} + (1-y(\tilde{x}_i,\tilde{x}_j))\frac{\partial \Delta_i(0,0)}{\partial \tilde{x}_i}) + \eta(\tilde{x}_i)(1-\beta_i)\frac{\partial \Delta_i(1,0)}{\partial \tilde{x}_i} > 0.$$

Hence endogenizing voter i's beliefs preserves optimality of the cutoff strategy for incompetent politician i.

Finally, differentiating with respect to $\tilde{x}_j$ yields

$$\frac{\partial \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)}{\partial \tilde{x}_j} = \frac{\partial \hat{I}(\tilde{x}_i, \tilde{x}_j; \Lambda_i)}{\partial \tilde{x}_j} + (1-\eta(\tilde{x}_i)-\eta(\tilde{x}_i)\beta_i)(y(\tilde{x}_i,\tilde{x}_j)\frac{\partial \Delta_i(0,1)}{\partial \tilde{x}_j} + (1-y(\tilde{x}_i,\tilde{x}_j))\frac{\partial \Delta_i(0,0)}{\partial \tilde{x}_j}) < 0.$$

Therefore, there exist cutoffs $(\tilde{x}_i^*, \tilde{x}_j^*)$ solving the system

$$\hat{I}^*(\tilde{x}_i^*, \tilde{x}_j^* \Lambda_i(\tilde{x}_i^*, \tilde{x}_j^*)) = 0 \quad \text{and} \quad \hat{I}^*(\tilde{x}_j^*, \tilde{x}_i^*, \Lambda_j(\tilde{x}_j^*, \tilde{x}_i^*)) = 0.$$

Analogous to Lemma A.9, the solution to the system is unique by the implicit function theorem by constructing the Jacobian matrix $\mathbf{J}$ and demonstrating that $|\mathbf{J}| > 0$. ■

Lemma A.11 *A unique pair of cutoffs $(\tilde{x}_i^*, \tilde{x}_j^*)$ exists, admitting a unique perfect Bayesian equilibrium to the international climate policy subgame. A politician of type θ in country i chooses policy $a_i = 1$ given signal x_i^θ with probability $\sigma^*(\theta_i, x_i^\theta) \in [0,1]$. These probabilities are*

$$\sigma^*(1, x_i^1) = x_i^1 = \omega.$$
$$\sigma^*(0, x_i^0) = 1 - G(\tilde{x}_i^*; \omega).$$

Upon observing policies a_i and a_j and signal s_i, the voter in country i reelects the politician with probability $F(\mu_i^*(a_i, s_i, a_j; \tilde{x}_i^*, \tilde{x}_j^*))$.

Proof Lemma A.11: Following any history in which politician i chooses policy a_i, voter i observes signal s_i, and politician j chooses policy a_j, the voter has posterior belief $P(\theta_i =$

$1|a_i, s_i, a_j) = \mu(a_i, s_i, a_j)$ as defined above, and reelects politician i iff $\mu(a_i, s_i, a_j) \geq \varepsilon_i$, which occurs with probability $F(\mu(a_i, s_i, a_j))$.

By Lemma A.8, the competent politician always follows her signal. By Lemma A.10, the incompetent politician plays a cutoff strategy such that she plays $a_i = 1$ iff $x_i^0 \geq \tilde{x}_i^*$, where $\tilde{x}_i^*$ exists and is a best response to both politician j's behavior and voter i's posterior beliefs. ■

Corollary A.5 *Politician i's cutoff $\tilde{x}_i^*$ is increasing in β_i.*

Proof of Corollary A.5: By the implicit function theorem,

$$\frac{d\tilde{x}_i^*}{d\beta_i} = -\frac{\begin{bmatrix} \frac{\partial \hat{I}^*(\tilde{x}_i,\tilde{x}_j,\Lambda_i)}{\partial \beta_i} & \frac{\partial \hat{I}^*(\tilde{x}_i,\tilde{x}_j,\Lambda_i)}{\partial k_j} \\ \frac{\partial \hat{I}^*(\tilde{x}_j,\tilde{x}_i,\Lambda_j)}{\partial \beta_i} & \frac{\partial \hat{I}^*(\tilde{x}_j,\tilde{x}_i,\Lambda_j)}{\partial k_j} \end{bmatrix}}{|\mathbf{J}|}.$$

It is clear that $\frac{\partial \hat{I}^*(\tilde{x}_j,\tilde{x}_i,\Lambda_j)}{\partial \beta_i} = 0$. Now differentiate $\hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)$ with respect to β_i to get

$$\frac{\partial \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)}{\partial \beta_i} = \eta(\tilde{x}_i)(y\Delta_i(0,1)+(1-y)\Delta_i(0,0)-\Delta_i(1,0))+(1-\eta(\tilde{x}_i)+\eta(\tilde{x}_i)\beta_i)(y\frac{\partial \Delta_i(0,1)}{\partial \beta_i}+(1-y)\frac{\partial \Delta_i(0,0)}{\partial \beta_i}).$$

Now, $\mu_i(1, 0, 1; \tilde{x}_i, \tilde{x}_j)$ is decreasing in β_i and $\mu_i(0, 0, 1; \tilde{x}_i, \tilde{x}_j)$ is increasing in β_i so $\Delta_i(0, 1)$ is decreasing in β_i. Similarly, $\mu_i(1, 0, 0; \tilde{x}_i, \tilde{x}_j)$ is decreasing in β_i and $\mu_i(0, 0, 0; \tilde{x}_i, \tilde{x}_j)$ is increasing in β_i so $\Delta_i(0, 0)$ is also decreasing in β_i. Hence $\frac{\partial \hat{I}^*(\tilde{x}_i,\tilde{x}_j,\Lambda_i)}{\partial \beta_i} < 0$ so the determinant of the matrix in the numerator is negative; by the implicit function theorem, $\frac{d\tilde{x}_i^*}{d\beta_i} > 0$. ■

Lemma A.12 *Fix β_j. Given equilibrium behavior in the international coordination game, special interest i's best response $\hat{\beta}_i(\beta_j) \in [0, 1]$ exists.*

Proof of Lemma A.12: Fix β_j (the proof is analogous for special interest j fixing β_i). Since β_j is a parameter, this proof is identical to Lemma A.2 albeit that the incompetent politician's cutoff is $\tilde{x}_i^*$ and not $\tilde{x}^*$ (although these cutoffs have the same relevant properties).

Special interest i has the objective function

$$\max_{\beta_i \in [0,1]} 1 - A_i\Big(\tilde{x}_i^*(\beta_i, \beta_j)\Big) - c(\beta_i).$$

Differentiating with respect to β_i yields the first-order condition

$$-\frac{dA_i}{d\beta_i} - c'(\beta_i) = 0,$$

and second-order condition

$$SOC = -\frac{d^2 A_i}{d\beta_i^2} - c''(\beta_i),$$

where $\frac{d^2 A_i}{d\beta_i^2} = -(1-\tau_i)\frac{d^2\tilde{x}_i^*}{d\beta_i^2}\Big(\pi g(\tilde{x}_i^*(\beta_i,\beta_j);1)+(1-\pi)g(\tilde{x}_i^*(\beta_i,\beta_j);0)\Big)-(1-\tau_i)(\frac{d\tilde{x}_i^*}{d\beta_i})^2\Big(\pi g'(\tilde{x}_i^*(\beta_i,\beta_j);1)+(1-\pi)g'(\tilde{x}_i^*(\beta_i,\beta_j);0)\Big)$. Similar to the argument made in the proof of Lemma A.2, a solution $\hat{\beta}_i(\beta_j)$ exists, either as the solution to the first-order condition or on the corner. ■

Lemma A.13 *Given equilibrium behavior in the international climate policy subgame, there exists an optimal pair* $(\beta_i^*, \beta_j^*) \in [0,1]^2$.

Proof of Lemma A.13: Since each interest group's best response is well-defined as shown in Lemma A.12, we now endogenize the behavior of the other interest group. Define the function $Z(\beta_i, \beta_j)$ as (analogous for group j):

$$Z(\beta_i,\beta_j) := (1-\tau_i)\frac{d\tilde{x}_i^*}{d\beta_i}\Big(\pi g(\tilde{x}_i^*(\beta_i,\beta_j);1) + (1-\pi)g(\tilde{x}_i^*(\beta_i,\beta_j);0)\Big) - c'(\beta_i),$$

where $\hat{\beta}_i(\beta_j)$ is the solution to $Z(\hat{\beta}_i,\beta_j) = 0$ for a fixed β_j, and $\frac{\partial Z(\hat{\beta}_i,\beta_j)}{\partial \beta_i} < 0$, by definition of $\hat{\beta}_i(\beta_j)$ being utility maximizing. The equilibrium levels of misreporting (β_i^*, β_j^*) are defined as the solution to the system

$$Z(\beta_i^*,\beta_j^*) = 0 \ \text{ and } \ Z(\beta_j^*,\beta_i^*) = 0.$$

To show that this system has a unique solution, define the Jacobian matrix as

$$\mathbf{J} = \begin{bmatrix} \frac{\partial Z(\beta_i^*,\beta_j^*)}{\partial \beta_i} & \frac{\partial Z(\beta_i^*,\beta_j^*)}{\partial \beta_j} \\ \frac{\partial Z(\beta_j^*,\beta_i^*)}{\partial \beta_i} & \frac{\partial Z(\beta_j^*,\beta_i^*)}{\partial \beta_j} \end{bmatrix}.$$

We know that $\frac{\partial Z(\beta_i^*,\beta_j^*)}{\partial \beta_i} < 0$ and $\frac{\partial Z(\beta_j^*,\beta_i^*)}{\partial \beta_j} < 0$. Differentiating with respect to β_j yields

$$\begin{aligned}\frac{\partial Z(\beta_i,\beta_j)}{\partial \beta_j} &= (1-\tau_i)\frac{d^2\tilde{x}_i^*}{d\beta_i d\beta_j}\Big(\pi g(\tilde{x}_i^*(\beta_i,\beta_j);1) + (1-\pi)g(\tilde{x}_i^*(\beta_i,\beta_j);0)\Big) \\ &\quad + (1-\tau_i)\frac{d\tilde{x}_i^*}{d\beta_i}\frac{d\tilde{x}_i^*}{d\beta_j}\Big(\pi g'(\tilde{x}_i^*(\beta_i,\beta_j);1) + (1-\pi)g'(\tilde{x}_i^*(\beta_i,\beta_j);0)\Big).\end{aligned}$$

While $\frac{\partial Z(\beta_i,\beta_j)}{\partial \beta_j}$ is not readily signed, observe that by symmetry, $sgn\frac{\partial Z(\beta_i,\beta_j)}{\partial \beta_j} = sgn\frac{\partial Z(\beta_j,\beta_i)}{\partial \beta_i}$. Then it is apparent that $|\mathbf{J}| \neq 0$, and hence nonsingular, so a solution to the system exists. ■

Proof of Proposition 2: Immediate from Lemmas A.11 and A.13. ■

International Cooperation: Results

This section proves Results 3, 4, 5, 6, and 7.

Proof of Result 3: Immediate from Lemma A.10 and the implicit function theorem,

$$\frac{d\tilde{x}_i^*}{d\tilde{x}_j} = -\frac{\partial \hat{I}^*(k_i, \tilde{x}_j, \Lambda_i)/\partial \tilde{x}_j}{\partial \hat{I}^*(k_i; \tilde{x}_j, \Lambda_i)/\partial k_i} > 0.$$

$\blacksquare$

Proof of Result 4: By the implicit function theorem,

$$\frac{d\tilde{x}_j^*}{d\beta_i} = -\frac{\begin{bmatrix} \frac{\partial \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)}{\partial \tilde{x}_i} & \frac{\partial \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)}{\partial \beta_i} \\ \frac{\partial \hat{I}^*(\tilde{x}_j, \tilde{x}_i, \Lambda_j)}{\partial \tilde{x}_i} & \frac{\partial \hat{I}^*(\tilde{x}_j, \tilde{x}_i, \Lambda_j)}{\partial \beta_i} \end{bmatrix}}{|\mathbf{J}|}.$$

It is clear that $\frac{\partial \hat{I}^*(\tilde{x}_j, \tilde{x}_i, \Lambda_j)}{\partial \beta_i} = 0$. From Corollary A.5, $\frac{\partial \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i)}{\partial \beta_i} < 0$ so the determinant of the matrix in the numerator is negative (as $\frac{\partial \hat{I}^*(\tilde{x}_j, \tilde{x}_i, \Lambda_j)}{\partial \tilde{x}_i} < 0$). Then $\frac{d\tilde{x}_j^*}{d\beta_i} > 0$. $\blacksquare$

Proof of Result 5: Country i's (j analogous) probability of climate action is

$$A_i(\tilde{x}_i^*) = \tau_i \pi + (1-\tau_i)\pi(1 - G(\tilde{x}_i^*; 1)) + (1-\tau_i)(1-\pi)(1 - G(\tilde{x}_i^*; 0)).$$

Differentiating with respect to β_i yields

$$\frac{dA_i}{d\beta_i} = -(1-\tau_i)\frac{d\tilde{x}_i^*}{d\beta_i}\Big(\pi g(\tilde{x}_i^*; 1) + (1-\pi) g(\tilde{x}_i^*; 0)\Big) < 0.$$
$$\frac{dA_j}{d\beta_i} = -(1-\tau_j)\frac{d\tilde{x}_j^*}{d\beta_i}\Big(\pi g(\tilde{x}_j^*; 1) + (1-\pi) g(\tilde{x}_j^*; 0)\Big) < 0.$$

Then, using the definitions of coordinated climate action, unilateral climate action, and coordinated climate inaction from the text, differentiating with respect to β_i yields

$$\text{coordinated climate action:} \quad A_i \frac{dA_j}{d\beta_i} + \frac{dA_i}{d\beta_i} A_j < 0.$$
$$\text{unilateral climate action:} \quad \frac{dA_i}{d\beta_i}(1 - 2A_j) + \frac{dA_j}{d\beta_i}(1 - 2A_i).$$
$$\text{coordinated climate inaction:} \quad -(1 - A_i)\frac{dA_j}{d\beta_i} - \frac{dA_i}{d\beta_i}(1 - A_j) > 0.$$

A sufficient condition that unilateral climate action is increasing in β_i is thus if $A_i > \frac{1}{2}$ and $A_j > \frac{1}{2}$. $\blacksquare$

Before proving Result 6, I prove a result about the cutoffs $\tilde{x}_i^*$ and $\tilde{x}_j^*$ (analogous to Lemma A.3).

Lemma A.14 *The following are true about the incompetent politician i's equilibrium cutoff (analogous for j):*

1. $\lim_{\pi \to 0} \tilde{x}_i^* = \infty$.
2. $\lim_{\pi \to 1} \tilde{x}_i^* = -\infty$.

Proof of Lemma A.14:

1. It is immediate that when $\pi \to 0$, we have $\eta(x) \to 0$ for any x, which also implies that $y(\tilde{x}_i, \tilde{x}_j) = (1-\tau_j)(1-G(\tilde{x}_j; 0))$. Then

   $$\lim_{\pi \to 0} \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i) = 2y(\tilde{x}_i, \tilde{x}_j) - 1 + y(\tilde{x}_i, \tilde{x}_j)\Delta_i(0,1; \tilde{x}_i, \tilde{x}_j) + (1 - y(\tilde{x}_i, \tilde{x}_j))\Delta_i(0,0; \tilde{x}_i, \tilde{x}_j).$$

   From the definition of the posterior beliefs induced by the cutoff, we know that $\Delta_i(0,1; \tilde{x}_i, \tilde{x}_j) < 0$ and $\Delta_i(0,0; \tilde{x}_i, \tilde{x}_j) < 0$. Hence $\lim_{\pi \to 0} \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i) < 0$, which means it is never optimal for the incompetent politician to choose $a = 1$, meaning $\tilde{x}_i^* \to \infty$.

2. It is immediate that when $\pi \to 1$, we have $\eta(x) \to 1$ for any x, which also implies that $y(\tilde{x}_i, \tilde{x}_j) = \tau_j + (1-\tau_j)(1-G(\tilde{x}_j; 1))$. Then

   $$\lim_{\pi \to 1} \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i) = 2y(\tilde{x}_i, \tilde{x}_j) - 1 + \beta_i y(\tilde{x}_i, \tilde{x}_j)\Delta_i(0,1; \tilde{x}_i, \tilde{x}_j) + \beta_i(1 - y(\tilde{x}_i, \tilde{x}_j))\Delta_i(0,0; \tilde{x}_i, \tilde{x}_j) + (1-\beta_i)\Delta_i(1,0; \tilde{x}_i, \tilde{x}_j).$$

   Now, given the posterior beliefs induced by the cutoffs, $\Delta_i(1,0; \tilde{x}_i, \tilde{x}_j)$ is constant in π, but $\Delta_i(0,1; \tilde{x}_i, \tilde{x}_j) > 0$ and $\Delta_i(0,0; \tilde{x}_i, \tilde{x}_j) > 0$. Hence $\lim_{\pi \to 1} \hat{I}^*(\tilde{x}_i, \tilde{x}_j, \Lambda_i) > 0$, which means it is always optimal for the incompetent politician to choose $a = 1$, meaning $\tilde{x}_i^* \to -\infty$.

■

Proof of Result 6: Recall that (β_i^*, β_j^*) is the solution to the system of equations

$$Z(\beta_i, \beta_j) = (1-\tau_i)\frac{d\tilde{x}_i^*}{d\beta_i}\Big(\pi g(\tilde{x}_i^*(\beta_i, \beta_j); 1) + (1-\pi) g(\tilde{x}_i^*(\beta_i, \beta_j); 0)\Big) - c'(\beta_i) = 0.$$
$$Z(\beta_j, \beta_i) = (1-\tau_j)\frac{d\tilde{x}_j^*}{d\beta_j}\Big(\pi g(\tilde{x}_j^*(\beta_j, \beta_i); 1) + (1-\pi) g(\tilde{x}_j^*(\beta_j, \beta_i); 0)\Big) - c'(\beta_j) = 0.$$

Observe that, by Lemma A.14, at $\pi = 0$ and $\pi = 1$, $Z(\beta_i, \beta_j) < 0$ for any $\beta_i > 0$ and any β_j, and $Z(\beta_j, \beta_i) < 0$ for any $\beta_j > 0$ and any β_i. Furthermore, $Z(\beta_i, \beta_j) = 0$ for $\beta_i = 0$ and $Z(\beta_j, \beta_i) = 0$ for $\beta_j = 0$. Hence $(\beta_i^*, \beta_j^*) = (0,0)$ is an equilibrium.

Further by Rolle's theorem there must be a $\hat{\pi}_i \in (0,1)$ where $\frac{\partial Z_(\beta_i,\beta_j)}{\partial \pi} = 0$, and there must be a $\hat{\pi}_j \in (0,1)$ where $\frac{\partial Z_(\beta_j,\beta_i)}{\partial \pi} = 0$, meaning that β_i^* and β_j^* are nonmonotonic in π. Note that it need not be the case that $\hat{\pi}_i = \hat{\pi}_j$.

Consider β_i^* (analogous for β_j^*). Partially differentiating yields

$$\begin{aligned}\frac{\partial Z(\beta_i,\beta_j)}{\partial \pi} = (1-\tau)\Big[& g(\tilde{x}_i^*;1)\frac{d\tilde{x}_i^*}{d\beta_i} + \pi g'(\tilde{x}_i^*;1)\frac{d\tilde{x}_i^*}{d\pi}\frac{d\tilde{x}_i^*}{d\beta} + \pi g(\tilde{x}_i^*;1)\frac{d^2\tilde{x}_i^*}{d\beta_i d\pi} \\ & - g(\tilde{x}_i^*;0)\frac{d\tilde{x}_i^*}{d\beta_i} + (1-\pi)g'(\tilde{x}_i^*;0)\frac{d\tilde{x}_i^*}{d\pi}\frac{d\tilde{x}_i^*}{d\beta_i} + (1-\pi)g(\tilde{x}_i^*;0)\frac{d^2\tilde{x}_i^*}{d\beta_i d\pi}\Big].\end{aligned}$$

Observe that at $\pi = 0$ and $\pi = 1$, $\frac{\partial Z(\beta_i,\beta_j)}{\partial \pi} = 0$, implying that such points are extrema, and we know that $\beta_i^* = 0$ in these cases. But because $\beta_i \in [0,1]$, these must be minima. Then the point $\hat{\pi}_i$ which is defined by Rolle's theorem must be an interior maximum such that β_i^* is increasing when $\pi < \hat{\pi}_i$ and decreasing when $\pi > \hat{\pi}_i$. Such a $\hat{\pi}_i$ is characterized by $\frac{\partial Z(\beta_i,\beta_j)}{d\pi} = 0$ and $\frac{\partial^2 Z(\beta_i,\beta_j)}{d\pi^2} \leq 0$. ■

Proof of Result 7: Given the definitions from the main text of $A_i(\tilde{x}_i^*)$ and $R_i(\tilde{x}_i^*)$:

$$R_i(\tilde{x}_i^*) - A_i(\tilde{x}_i^*) = \tau_i(1-\pi) + (1-\tau_i)(1-\pi)(2G(\tilde{x}_i^*;0)-1).$$

Differentiating with respect to β_i and β_j yields

$$\begin{aligned}\frac{d(R_i(\tilde{x}_i^*) - A_i(\tilde{x}_i^*))}{d\beta_i} &= 2(1-\tau_i)(1-\pi)g(\tilde{x}_i^*;0)\frac{d\tilde{x}_i^*}{d\beta_i} > 0. \\ \frac{d(R_i(\tilde{x}_i^*) - A_i(\tilde{x}_i^*))}{d\beta_j} &= 2(1-\tau_i)(1-\pi)g(\tilde{x}_i^*;0)\frac{d\tilde{x}_i^*}{d\tilde{x}_j^*}\frac{d\tilde{x}_j^*}{d\beta_i} > 0.\end{aligned}$$

■

B Additional Descriptive Figures

Figure A.1 demonstrates that the lion's share of these laws are related to climate mitigation. The growth of mitigation laws over time is notable, especially as nearly all NDCs under the Paris framework include mitigation measures. Moreover, as mitigation is nationally costly but provides global benefits, theories of collective action would predict a stagnation or underprovision of mitigation laws. Other laws aim to address adaptation, disaster risk management, and loss and damages. These policies, while much more difficult to measure, are also increasing in frequency as climate change's effects become more pronounced, especially in the Global South.

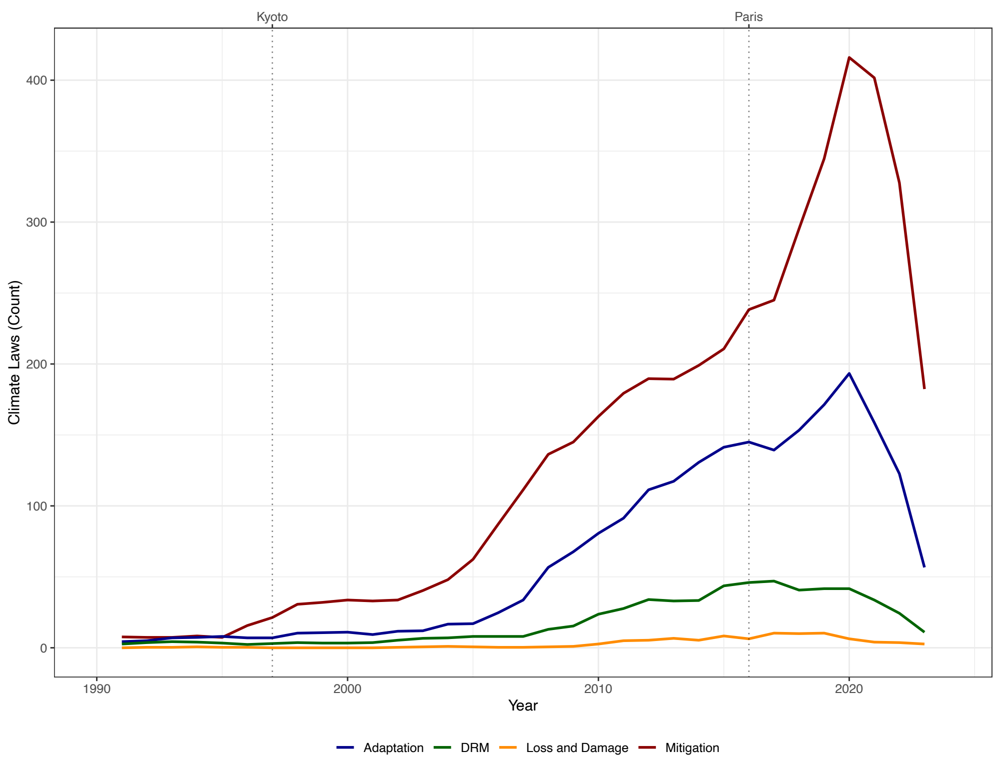

Figure A.1: Climate Laws by Type

Figure A.2 shows the correlation in the OECD Environmental Policy Stringency Index across countries. The figure shows that all cross-country correlations of environmental policy stringency are positive, and almost all of the are statistically distinguishable from zero (those that are not are in red). These positive correlations are meaningful because they suggest that increases in climate ambitions across countries are complementary, not substitutable, as theories based in free-riding would predict.

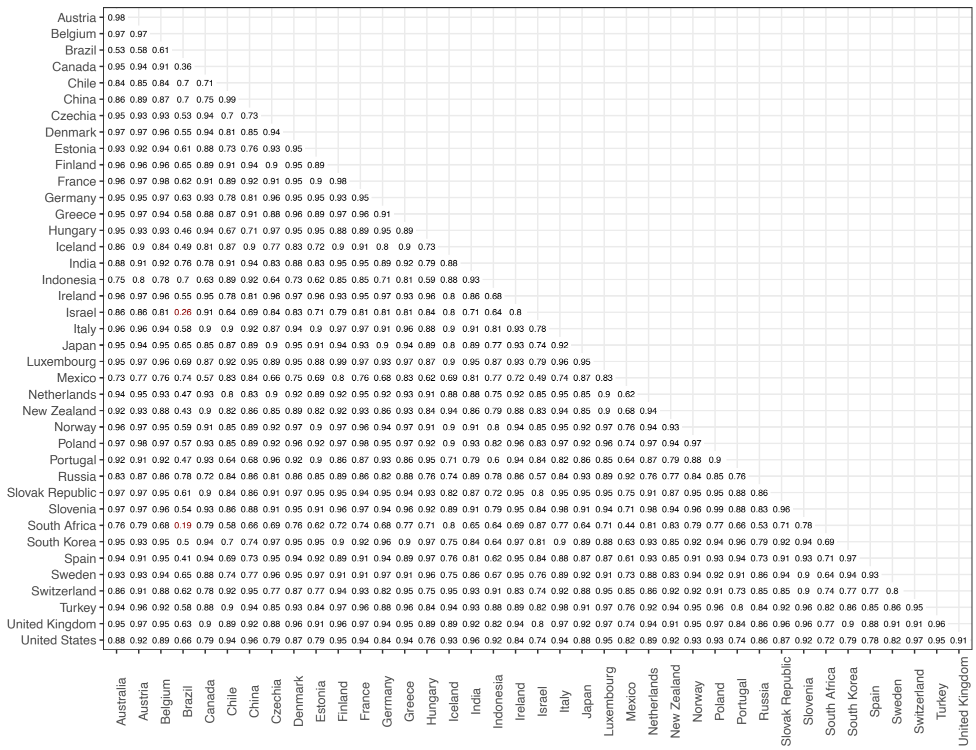

Figure A.2: Correlation Between Countries' Environmental Policy Stringency, 1990-2020

C Exxon Source Documents

- Exxon scientist James Black's 1978 memo notes the scientific consensus that the climate is affected by fossil fuels. Notably, he writes that "present thinking holds that man has a time window of five to ten years before the need for hard decisions regarding changes in energy strategies might become critical." The document is available at https://climateintegrity.org/uploads/deception/1978-Exxon-BlackMemo.pdf.

- Exxon scientist Roger Cohen's 1982 memo summarizes findings about Exxon's internal climate modeling. It documents the projected relationship between increased carbon dioxide in the atmosphere and changes in the Earth's climate. It further discusses the scientific consensus around this result. The document is available at https://www.climatefiles.com/exxonmobil/1982-exxon-memo-summarizing-climate-modeling-and-co2-greenhouse-effect-research/.

- In 1996 and 1998, Exxon released pamphlets to the masses that sought to inject doubt into the public discourse about the validity of climate science and the subsequent need for policy action. In particular, one such pamphlet was entitled "Global climate change: everyone's debate," and is available at https://www.climatefiles.com/exxonmobil/1998-exxon-pamphlet-global-climate-change-everyones-debate/. The pamphlet "Global warming: who's right?" admonishes readers not to "ignore the facts" about climate change and is available at https://climateintegrity.org/uploads/deception/1996-Exxon-Global-Warming-Whos-Right.pdf.

- The "Victory Memo" of 1998 makes the goal to inject uncertainty into the public sphere clear: "victory will be achieved when average citizens 'understand' (recognize) uncertainties in climate science," and "recognition of uncertainty becomes part of the 'conventional wisdom."' The memo can be found at https://www.climatefiles.com/trade-group/american-petroleum-institute/1998-global-climate-science-communications-team-action-plan/.

- "The Greenhouse Effect" is a report published by a working group of Shell scientists in 1988 documents potential climate impacts, including rising sea levels, ocean acidification, and human migration, from continued fossil fuel production. The full document is available at https://www.documentcloud.org/documents/4411090-Document3.html.

- The Global Climate Coalition was a lobbying group of several large oil and gas companies that operated between 1989 an 2001. Its primary function was to coordinate messaging against global climate action like the ratification of the Kyoto Protocol. In 1995, the GCC internally circulated *Predicting Future Climate Change: A Primer*, which summarized the state of climate science. Notably, it reads, "The scientific basis for the Greenhouse Effect and the potential impact of human emissions of greenhouse gases such as CO2 on climate is well established and cannot be denied." The full document is available at https://www.ucsusa.org/sites/default/files/attach/2015/07/Climate-Deception-Dossier-7_GCC-Climate-Primer.pdf.

- While the GCC internally circulated *Predicting Future Climate Change: A Primer*, its public-facing publications of the time were very different. In 1995, it also published “Climate Change: Your Passport To The Facts,” a booklet allegedly intended to introduce readers to essential facts about climate change. Facts include that “the notion that scientists have reached consensus that man-made emissions of greenhouse gases are leading to a dangerous level of global warming is not true” and “computer climate models, which are the basis for ”predictions” of global climate change, suffer from severe flaws.” The document is available at https://www.worthingtoncaron.com/documents/1995-CLIMATE-CHANGE-YOUR-PASSPORT.pdf.

- ExxonMobil published a series of newspaper ads in order to sow doubt into the public about climate science. In the spring of 2000, ExxonMobil rad the ad “Unsettled Science” in major news outlets (e.g., the *New York Times*). These ads also tried to discredit climate scientists. Scientists like Lloyd Keigwin later responded in the *Wall Street Journal* complaining that ExxonMobil had distorted his work by suggesting it supported the notion that global warming was just a natural cycle.[1]

[1] https://insideclimatenews.org/news/22102015/exxon-sowed-doubt-about-climate-science-for-decades-by-stressing-uncertainty/