**Appendix**

The following contains the proofs for the propositions and corollaries of the formal model and a description of the empirical estimation approach.

*Proof of Proposition 1: Existence of Credible Commitments Equilibrium*

Where necessary, I index the optimal initial and final policies chosen by bad governments with the subscript : and . For good governments, I use the subscript. Where there is no need to distinguish between government types, I omit the subscripts.

For the audience to choose , it must be the case that . Rewriting the audience's expected utilities:

where .

For the audience to choose , it must be the case that . As above, the audience's expected utilities are:

where .

Derivations of and , as well as optimal policies chosen by good governments and dispute probabilities are shown in the proofs for subsequent propositions.

*Proof of Proposition 2: Optimal Post-mobilization policy*

After mobilization, the home government faces the following optimization problem:

The proof follows from rearranging the first order conditions of the post-mobilization maximization problem, .

The ratio of the audience and home government's marginal utilities matches the (inverse) ratio of their strength after mobilization. If the home government and audience's utility functions, and , were identical apart from their maximization points and were symmetrical, then the optimal policy would be an α-weighted combination of the two ideal points, . For instance, this would be the case if both the home government and audience held preferences represented by the often-used quadratic loss function. If the audience and the home government share the same ideal point, , as in the case of a ``good'' government, then .

*Proof of Proposition 3 and 4: Probability of a Dispute and Optimal Initial Policy*

Before describing optimal initial policy, I describe the probability of a dispute. The utility to the foreign government of initiating a dispute is , and the utility of not doing so is . In a CCE, the foreign government initiates a dispute if and only if their costs are lower than their expected gains:

Recall, for a good home government, , and for a bad home government, . For a good home government, therefore, the foreign government only initiates a dispute if it draws a negative litigation costs, i.e. it has some extraneous benefit to initiating a dispute, apart from the potential effects on home's policies. Facing a bad home government, the benefit of a dispute comes from the effect that any subsequent audience mobilization will have on changing the initial tariff policy to a new, lower final policy. If the foreign government draws a litigation cost that is higher than the benefits from changing the home government's policy, then it will not initiate a dispute. The probability of a dispute for a particular initial policy, which I call , is the probability that the foreign government draws a low enough litigation cost that it will choose to initiate a dispute.

The home government's initial optimization problem and related first order condition are:

For a good home government, their optimal policy choice is . Good home governments can do no better by choosing a different initial policy. If the foreign government draws a negative litigation cost and initiates a dispute, then the good home government will still choose . If the foreign government draws a higher litigation cost, they will not initiate a dispute and the audience will not mobilize, leaving the home government's ideal policy in place.

Observe that for bad governments, . The home government can do no better by choosing an initial policy higher than , such that . Lowering the policy to decreases the probability of a dispute and leaves the home government better off if they avoid a dispute. Similarly, the home government can do no better by choosing a policy lower than , such that . Raising the policy to lowers the probability of a dispute by decreasing the distance between and and leaves the home government better off if they avoid a dispute.

Rewriting the FOC for the home government's maximization problem associated with yields:

Since is uninfluenced by , we can rewrite the FOC as:

where is the total derivative of the FOC with respect to and is the total derivative of the FOC with respect to .

Rearranging yields:

Substituting in the total derivatives, and yields:

Since for the uniform distribution, this equation can be signed by observing that and for all . It follows that . This implies that ``inherits'' the properties of that are described in Corollary 1.

*Proof of Proposition 5: Audience Effects on Optimal Initial Policy*

This proof builds off of the proof for Proposition 4 which showed that . Now, we consider whether . If , then equilibrium increases in result in *smaller* accompanying increases in . Since is distributed uniformly, this would imply that the post-dispute effect dominates.

Recall the expression for with the uniform distribution simplifies to:

Since Proposition 4 implies that the numerator and denominator have the same sign (+), for it must be the case that:

yielding the condition stated in Proposition 5.

*Details of Empirical Model*

Following Imai and Van Dyk (2005), I let the observed multinomial variable, , take on a distinct value depending on the status of tariff *i* at time *t*. Let *j = 1,2,3* index the 3 statuses, *WTO Dispute, Unilateral Removal, In Effect*. Call *j=3*, *In Effect*, the base category. Let be a vector of 2 latent variables, associated with *WTO Dispute* and *Unilateral Removal*, for tariff *i* at time *t*. The observed variable, is modeled in terms of via:

where represents the largest value in the vector . The latent variables are modeled as a function of the *k* observed covariates.

 is a *2 x k* matrix of observed covariates and is a *k x 1* vector of coefficients. is a positive definite *2 x 2* matrix. For identification, the model assumes that . The Bayesian approach implemented here uses the MCMC procedure developed by Imai and Van Dyk (2005) to sample to sample from posterior distributions of and , based on particular prior distributions. I used very agnostic priors, where each element of is distributed normally with mean 0 and variance 100. Setting the prior variance to 100 means that the prior distribution is very diffuse and unlikely to influence results. For the main MNP model, I used a burn-in of 20,000 draws and kept every fourth draw from 70,000 subsequent draws. For the models with calendar month and age polynomials included as covariates, I set the prior variance to 80, used a 15,000 draw burn-in, and kept every fourth draw from 60,000 subsequent draws.