Online Supporting Information
Signaling with Reform: How the Threat of Corruption Prevents Informed Policymaking

Keith E. Schnakenberg* Ian R. Turner†

Contents

*Assistant Professor of Political Science, Washington University in St. Louis. Contact: keschnak@wustl.edu.
†Assistant Professor of Political Science, Yale University. Contact: ian.turner@yale.edu.

A Extensions

This appendix provides formal analysis that forms the basis of informal discussions in the main body of the paper. First, we relax the assumptions that $\pi_A = \pi_B = \pi$ and that candidates play symmetric strategies. Second, we characterize the conditions under which interest groups benefit from self-regulating so that they can not bribe corrupt politicians.

A.1 Asymmetric corruption

In this section we relax the assumption that each candidate is equally likely to be sincere. We prove analogous results to those presented in the main body of the paper. The main difference is that we relax our focus on symmetric strategies to mirror our relaxation of model symmetry.

Suppose that politician A is more likely to be corrupt than politician B: $\pi_A < \pi_B$. This means that candidate B has an ex ante electoral advantage. That is, in the absence of new information the voter retains her prior beliefs that A and B are sincere/corrupt and therefore elects politician B in that case (as opposed to each candidate being elected with equal probability in the symmetric corruption model). Note also that the policymaking strategies of winning politicians and interest group bribery/lobbying strategies are equivalent because at that point of the game politician type is revealed to the group. So nothing changes from Proposition 1 in the baseline symmetric corruption model presented in text: corrupt politicians implement $x = 1$ if $b = 1$ and $x = 0$ otherwise, sincere politicians implement $x = 1$ when $m = 1$ and $x = 0$ when $m = 0$, and the interest group always bribes corrupt politicians and lobbies sincere politicians only when $\theta = 1$. Moreover, it is still optimal for the voter to simply elect the candidate most likely to be sincere and elect either candidate with equal probability when they are both equally likely to be sincere.

The results do change when analyzing the candidate access decisions. We proceed in a similar manner from the analysis of the symmetric corruption model presented in the main body of the paper. Notice first that Lemma 2 still holds in this setting. Corrupt candidates have no incentive to ban access by the argument in the proof of Lemma 2. This is true regardless of the asymmetry between π_A and π_B since both are still positive and less than one. We proceed by establishing the conditions for a separating reform equilibrium, a pooling access equilibrium, and equilibria in asymmetric strategies in which one candidate pools on access and one separates with reform.

Separating reform equilibrium. The following result provides the conditions required to support a reform equilibrium when $\pi_A < \pi_B$.

Proposition 4. *Suppose $\pi_A < \pi_B$. The conditions to support a separating reform equilibrium are the same as in Proposition 2 (i.e., when $\pi_A = \pi_B = \pi$).*

Proof of Proposition 4. Notice first that the argument in Lemma 2 implies that corrupt candidates always grant access. So we need to show the conditions that support sincere candidates banning access. If candidate A is sincere and plays the posited strategy (banning access) then he wins with probability $\frac{1}{2}$ when B is sincere since both play the same separating strategy and the voter correctly believes both to be sincere. If B is corrupt then A wins for sure. This yields the following expected utility for banning access,

$$\begin{aligned} EU_A(a_A = 0|\tau_A = S, \pi_B) &= -\pi_B\left(\frac{1}{2}(1-q) + \frac{1}{2}(1-q)\right) - (1-\pi_B)(1-q), \\ &= -(1-q). \end{aligned}$$

In contrast, if A deviates to $a = 1$ then he loses for sure when B is sincere since the voter believes he is corrupt and wins with probability one-half if B is corrupt since the voter believes both are corrupt.

$$\begin{aligned} EU_A(a_A = 1|\tau_A = S, \pi_B) &= -\pi_B\left((1-q)\right) - (1-\pi_B)\left(\frac{1}{2}(0) + \frac{1}{2}q\right), \\ &= -\pi_B(1-q) - \frac{1}{2}(1-\pi_B)q. \end{aligned}$$

This yields the following incentive compatibility condition for candidate A to continue to ban access when sincere (given B does the same):

$$-(1-q) > -\pi_B(1-q) - \frac{1}{2}(1-\pi_B)q,$$

which is satisfied for all $\pi_B \in (0,1)$ when $q \in \left(\frac{2}{3}, 1\right)$.

Again, expected utility calculations for candidate B are exactly the same once we substitute in π_A for π_B. This is because we are assuming that both candidates play symmetric strategies in this particular equilibrium even though the probabilities of being corrupt are asymmetric across candidates.

$$\begin{aligned} EU_B(a_B = 0 | \tau_B = S, \pi_A) &= -\pi_A \left(\frac{1}{2}(1-q) + \frac{1}{2}(1-q)\right) - (1-\pi_A)(1-q), \\ &= -(1-q). \end{aligned}$$

$$\begin{aligned} EU_B(a_B = 1 | \tau_B = S, \pi_A) &= -\pi_A (1-q) - (1-\pi_A) \left(\frac{1}{2}(0) + \frac{1}{2}q\right), \\ &= -\pi_A(1-q) - \frac{1}{2}(1-\pi_A)q. \end{aligned}$$

$$a_B^*(\tau_B = S, \pi_A) = 0 \iff -(1-q) > -\pi_A(1-q) - \frac{1}{2}(1-\pi_A)q.$$

Thus, for all $\pi_A, \pi_B \in (0,1)$ we can support a separating equilibrium where sincere candidates ban access and corrupt candidates grant access, the voter learns candidate types with certainty, and elects the candidate identified as sincere or elects either candidate with equal probability when both are of the same type so long as $q \in \left(\frac{2}{3}, 1\right)$. This is the same condition as in the case in which both candidates are sincere with equal probability: $q > q^{\text{Reform}}(\pi)$. ■

Pooling access equilibrium. The following result provides the conditions required to support an access equilibrium when $\pi_A < \pi_B$. In this case the conditions to support a pooling equilibrium in which all candidates grant interest group access regardless of type are more demanding.

Proposition 5. *Suppose $\pi_A < \pi_B$. Then the conditions to support an access equilibrium are more demanding than when $\pi_A = \pi_B = \pi$. Specifically, instead of $q \in \left(\frac{1}{2}, \frac{2}{3-\pi}\right)$, the relevant condition is $q \in \left(\frac{1}{2}, \frac{1}{2-\pi_B}\right)$ for all $\pi_B \in (0,1)$.*

Proof of Proposition 5. A is ex ante disadvantaged: $\pi_A < \pi_B$. This implies that candidate B wins the election for sure when both candidates pool on granting access because the voter is not able to learn anything about candidate types through access decisions, retains her prior about each candidate, and elects candidate B since he is ex ante more likely to be sincere. Lemma 2 shows that corrupt candidates always want to grant access so we show the conditions for sincere candidates to also grant access.

Candidate B's expected utility for granting access given that A is also granting access regardless of type is given by,

$$\begin{aligned} EU_B(a_B^* = 1 | \tau_B = S, \pi_A) &= -\pi_A(0) - (1-\pi_A)(0), \\ &= 0. \end{aligned}$$

B wins the election and, because he granted access, learns θ, implements policy accordingly and losing nothing in utility. His expected utility for deviating to $a = 0$ is given by (assuming that the voter believes deviations of this sort signal sincere type),

$$\begin{aligned} EU_B(a = 0 | \tau_B = S, \pi_A) &= -\pi_A(1-q) - (1-\pi_A)(1-q), \\ &= -(1-q). \end{aligned}$$

In this case, B still wins the election for sure,[1] but now because interest group access was banned does not receive information regarding θ, implements $x = 0$, and in expectation loses one with probability $(1-q)$. Obviously in this case politician B always wants to pool on $a = 1$ since $0 > -(1-q)$ for all $q \in \left(\frac{1}{2}, 1\right)$.

Now consider the incentives for candidate A. A's expected utility for continuing to pool on

[1]The same Intuitive Criterion argument in the proof of Proposition 2 applies here: the voter believes that this deviation identifies the candidate as sincere.

$a = 1$ is given by,

$$\begin{aligned} EU_A(a_A^* = 1 | \tau_A = S, \pi_B) &= -\pi_B(0) - (1-\pi_B)(q), \\ &= -q(1-\pi_B). \end{aligned}$$

In this case, A always loses the election, but if B is a sincere type A loses nothing from a policy perspective since B matches policy to the state. However, when B is corrupt (with probability $1-\pi_B$) A expects to lose on policy with probability q since B will always implement $x = 1$. If A deviates to $a = 0$, and the voter accordingly updates that A is sincere and therefore A will win (again this is the only off-path belief that satisfies the Intuitive Criterion as in Proposition 2), he receives the following expected utility,

$$\begin{aligned} EU_A(a_A = 0 | \tau_A = S, \pi_B) &= -\pi_B(1-q) - (1-\pi_B)(1-q), \\ &= -(1-q). \end{aligned}$$

In this case candidate A can win the election, but this comes at the cost of information once he has won since he had to ban group access to do so. Therefore, he implements $x = 0$ since $q > \frac{1}{2}$ and expects to lose on policy with probability $1-q$. Combining these two expected utility expresses yields the incentive compatibility condition for A to continue to pool on access:

$$-q(1-\pi_B) > -(1-q),$$

which is satisfied for all $\pi_B \in (0,1)$ so long as $q \in \left(\frac{1}{2}, \frac{1}{2-\pi_B}\right)$.

Now, the upper bound has changed from the case of symmetric corruption pooling. In that case, $q < \frac{2}{3-\pi}$ supported pooling and in this case $q < \frac{1}{2-\pi_B}$ is (necessary and) sufficient. Obviously, $\frac{2}{3-\pi} > \frac{1}{2-\pi_B}$, which highlights the fact that the conditions on q to support the access equilibrium are more demanding when $\pi_A < \pi_B$. In both cases this upper bound is increasing in π_i. Also notice that since in this case B always wants to pool when A does that π_A makes no difference (it does not restrict the range of parameters in which this pooling on access behavior is an equilibrium), so π_B is the relevant corruption probability due to how it restricts A's behavior. That is, only the probability of B being corrupt is relevant to support pooling since A is the politician with the incentive to deviate to a separating strategy. ■

Asymmetric strategy equilibria. The following result characterizes the conditions under which the two candidates play different strategies. That is, one candidate pools on access while the other separates by instituting reform when sincere. We state and prove the result without reference to particular candidate identity because the result holds for any ordering of candidate identity and corruption probabilities by substituting A or B for i or j. Notice that this has to do with relaxation of the symmetric strategies assumption in the model presented in the main body of the paper. This result does not depend on whether probabilities of corruption are equal or different. Thus, this result would also hold in the main analysis.

Proposition 6. *Suppose candidates can play asymmetric strategies. Then when $q \in \left(\frac{1}{2-\pi_j}, 1\right)$ we can support an equilibrium in which candidate i separates (as in a reform equilibrium) and candidate j pools (as in an access equilibrium), for all $i \neq j$.*

Proof of Proposition 6. Lemma 2 implies that corrupt candidates always grant access so we focus on the incentives for sincere candidates. Suppose first that candidate i separates by choosing $a_i^*(\tau_i) = 0$ when $\tau_i = S$ and $a_A^*(\tau_i) = 1$ when $\tau_i = C$. Further, suppose that candidate j pools on access so that $a_B^*(\tau_j) = 1$ for all $\tau_j \in \{S, C\}$. The voter best responds by electing candidate i following observation of $a_i^* = 0$ and electing candidate j if both candidates grant access since in this case the voter correctly believes candidate i is corrupt while candidate j, by virtue of pooling, is sincere with probability $\pi_j > 0$ (i.e., the voter's prior that j is sincere). If both i and j choose to ban access then the voter elects either with equal probability.[2]

[2]The Intuitive Criterion argument in the proof of Proposition 2 implies that if j deviates and bans access then the voter places full mass on j being a sincere type.

First, consider candidate i's expected utility for banning access when he is sincere:

$$\begin{aligned} EU_i(a_i^* = 0 | \tau_i = S, \pi_j) &= -\pi_j(1-q) - (1-\pi_j)(1-q), \\ &= -(1-q). \end{aligned}$$

In this case i always wins the election since j is pooling on access and therefore i is more likely to be sincere from the voter's perspective. However, since i banned access he does not receive any information about θ from the group, implements $x = 0$, and mismatches policy and the state with probability $1-q$. In contrast, candidate i's expected utility for deviating to granting access is given by,

$$\begin{aligned} EU_i(a_i^* = 1 | \tau_i = S, \pi_j) &= -\pi_j(0) - (1-\pi_j)(q), \\ &= -(1-\pi_j)q. \end{aligned}$$

In this case i loses the election for sure because the voter infers he is corrupt. He loses nothing on policy if candidate j is sincere, since in that case j matches policy to the state. If instead candidate j is corrupt, then the group bribes j and he implements $x = 1$ for sure, which in expectation leads to a policy loss with probability q. Thus, candidate i will continue to separate when candidate j pools if,

$$-(1-q) > -(1-\pi_j)q,$$

which is satisfied for all $\pi_j \in (0,1)$ so long as $q \in \left(\frac{1}{2-\pi_j}, 1\right)$.

Now, given that candidate i is separating what are the conditions that support candidate j's pooling on access? First, consider j's expected utility when sincere of granting access given that i is separating:

$$\begin{aligned} EU_j(a_j^* = 1 | \tau_j = S, \pi_i) &= -\pi_i(1-q) - (1-\pi_i)(0), \\ &= -\pi_i(1-q). \end{aligned}$$

In this case, if i is sincere j will lose and have to eat the probability that i mismatches policy to the state given that he will get no further information from the group since access was banned. If i is corrupt then j wins and will match policy to the state thereby losing zero in utility. In contrast, if j deviates and signals $a_j = 0$ his expected utility is,

$$\begin{aligned} EU_j(a_j = 0|\tau_j = S, \pi_i) &= -\pi_i\left(\frac{1}{2}(1-q)+\frac{1}{2}(1-q)\right)-(1-\pi_i)(1-q), \\ &= -(1-q). \end{aligned}$$

The voter updates that j is sincere (since this is the only off-path belief that survives the Intuitive Criterion) and therefore elects i and j with equal probability when i is sincere and also bans access. In this case whoever wins will mismatch policy to the state with probability $1-q$. If i is corrupt and grants access then j will win for sure but will not learn anything about θ, implement $x = 0$, and this will lead to a loss of one with probability $1-q$. Thus, when j is sincere he always wants to stick to pooling on access given that i is separating since $-\pi_i(1-q) > -(1-q)$ for all $\pi_i \in (0,1)$.

Overall, we have an equilibrium in which i separates with access decisions and j pools on access any time that $q > \frac{1}{2-\pi_j}$, as stated in the result. ■

A.2 Interest group self-regulation

In this section we explore whether and when the interest group may benefit from self-regulation. That is, when will the interest group benefit from committing ex ante to not bribing corrupt politicians? We explore this question from the perspective of interest group ex ante welfare.

Suppose that the interest group has committed to no longer bribe corrupt politicians that win office, but it can still engage in substantive lobbying. So $b = 0$ always. Nothing in the policymaking stage of the game changes *except* that the interest group can no longer bribe corrupt politicians that have won office. Thus, corrupt politicians always implement $x = 0$ since $b = 0$, sincere politicians always implement $x = 1$ when $m = 1$ and $x = 0$ when $m = 0$ because the interest group still prefers to separate by lobbying $m = 1$ if and only if $\theta = 1$ (which implies $m = 0$ when $\theta = 0$). Similarly, the voter's voting strategy still does not change: she votes for the candidate most likely to be sincere and elects either candidate with equal probability when each candidate is equally likely to be sincere.

To begin the analysis we first show that the interest group never benefits from committing to no bribery when the candidates play access equilibrium strategies.

Lemma 3. *The interest group never benefits from self-regulating by committing to no bribery when candidates play access equilibrium strategies.*

Proof of Lemma 3. First, consider the interest group's welfare when there is no bribery and all candidates grant access. In this case each candidate wins the election with equal probability since the voter cannot differentiate candidate types. If the group cannot bribe corrupt politicians then its ex ante expected welfare in a pooling access equilibrium is given by,

$$\begin{aligned} W_G(\text{No bribes}|\text{Access}) &= \pi(q(0)+(1-q)(1-\alpha_1))+(1-\pi)(0), \\ &= \pi((1-q)(1-\alpha_1)). \end{aligned}$$

With probability π the winning candidate is sincere. In this case the group receives zero if $\theta = 0$ since it will not lobby, $m = 0$, and the candidate will implement $x = 0$. This occurs with probability q. With probability $1-q$, $\theta = 1$, the group will lobby $m = 1$, and the politician will implement $x = 1$. This yields an expected payoff of $(1-\alpha_1)$. With probability $1-\pi$ the winner is corrupt, but since bribery has been banned the group can not affect the politician's implementation of $x = 0$, which yields a payoff of zero. Compare this with the interest group's welfare in the pooling access equilibrium when bribing corrupt candidates is possible:

$$\begin{aligned} W_G(\text{Bribes}|\text{Access}) &= \pi(q(0)+(1-q)(1-\alpha_1))+(1-\pi)(1-\kappa), \\ &= \pi((1-q)(1-\alpha_1))+(1-\pi)(1-\kappa). \end{aligned}$$

The group's expected payoffs when a sincere candidate wins are the same as above. When the winning candidate is corrupt the group pays a bribe $b = 1$ at cost κ, the candidate implements

$x = 1$, and the group receives $1-\kappa$. This last component of group welfare is the difference between bribery and no bribery. That is, the net welfare from the interest group's perspective when bribery is banned is given by,

$$\begin{aligned}
W_G(\text{No bribes}|\text{Access}) - W_G(\text{Bribes}|\text{Access}) &= \pi((1-q)(1-\alpha_1)) - \pi((1-q)(1-\alpha_1)) \\
&\quad -(1-\pi)(1-\kappa), \\
&= -(1-\pi)(1-\kappa).
\end{aligned}$$

The group derives a net benefit from being able to bribe corrupt politicians equal to $(1-\pi)(1-\kappa)$ relative to not being able to bribe when all candidates are granting access. Thus, the group always prefers to retain its ability to bribe when candidates will play access equilibrium strategies for sure (e.g., when $q < \frac{2}{3}$). ■

Next, we establish that when the interest group has self-regulated by committing to no bribery sincere candidates no longer have incentives to separate by banning access.

Lemma 4. *Suppose that the interest group has committed to no bribery. Then all candidates grant interest group access.*

Proof of Lemma 4. Corrupt candidates continue to grant group access by the argument in Lemma 2. However, the incentives for sincere candidates to separate by banning access have changed. Consider a sincere candidate's expected payoff for banning access, given that corrupt candidates grant access:

$$\begin{aligned}
EU_i(a_i = 0|\tau_i = S, a_{-i}, \pi) &= -\pi\left(\frac{1}{2}(1-q) + \frac{1}{2}(1-q)\right) - (1-\pi)(1-q), \\
&= -(1-q).
\end{aligned}$$

If candidate i faces another sincere candidate then no matter who wins $x = 0$ is implemented and fails to match the state with probability $1-q$. Similarly, if i faces a corrupt candidate then he wins for sure, but since access was banned implements $x = 0$ and fails to match policy to the state with probability $1-q$. In contrast, if a sincere candidate i deviates to granting access then his expected payoff is given by,

$$\begin{aligned} EU_i(a_i=1|\tau_i=S,a_{-i},\pi) &= -\pi(1-q)-(1-\pi)\left(\frac{1}{2}(0)+\frac{1}{2}(1-q)\right), \\ &= -\pi(1-q)-\frac{1}{2}(1-\pi)(1-q). \end{aligned}$$

With probability π the sincere candidate loses for sure because he is facing another sincere candidate (who is still separating) and receives an expected payoff of $-(1-q)$. With probability $1-\pi$ the other candidate is corrupt and the sincere candidate wins half of the time and gets to match policy to the state, but half the time he loses and because bribery is banned the corrupt winner implements $x=0$, which yields an expected payoff of $-(1-q)$. We can no longer support sincere candidates optimally separating however since,

$$-(1-q) < -\pi(1-q)-\frac{1}{2}(1-\pi)(1-q),$$

for all $q \in \left(\frac{1}{2},1\right)$ and $\pi \in (0,1)$. Thus, now that the group committed to no bribery sincere candidates will no longer separate by banning interest group access.

To complete the proof we show that when the interest group cannot bribe sincere politicians prefer to pool on granting access. If a sincere candidate who is granting access faces another sincere candidate also granting access then each win with probability one-half, but no matter who wins policy will ultimately match the state yielding zero policy loss. If a sincere candidate granting access faces a corrupt candidate granting access then each win with probability one-half. If the sincere candidate wins he matches policy to the state. If the corrupt candidate wins he implements $x=0$ since there is no bribery. This will fail to match the state with probability $1-q$. The sincere candidate's expected payoff for pooling on access is then,

$$\begin{aligned} EU_i(a_i = 1|\tau_i = S, a_{-i}, \pi) &= -\pi\left(\frac{1}{2}(0) + \frac{1}{2}(0)\right) - (1-\pi)\left(\frac{1}{2}(0) + \frac{1}{2}(1-q)\right), \\ &= -\frac{1}{2}(1-\pi)(1-q). \end{aligned}$$

A deviation to banning access leads the sincere candidate to win with certainty regardless of his opponents type, but because he banned access he always implements $x = 0$ which fails to match the state with probability $1-q$. This yields an expected payoff of $EU_i(a_i = 0|\tau_i = S, a_{-i}, \pi) = -(1-q)$. Thus, sincere candidates will always grant access so long as $-\frac{1}{2}(1-\pi)(1-q) > -(1-q)$, which is satisfied for all $q \in \left(\frac{1}{2}, 1\right)$ and $\pi \in (0,1)$. ■

Now suppose that we are in an environment in which candidates play a separating reform equilibrium. The following result characterizes when the interest group benefits from committing ex ante to not bribing corrupt politicians that win office. Lemma 4 implies that in that case the candidates instead play a pooling access equilibrium. Thus, the trade-off for the interest group is between continuing to be able to bribe corrupt candidates but having sincere candidates identify themselves by banning access and self-regulating so they cannot bribe corrupt winners but inducing sincere candidates to grant them lobbying access.

Proposition 7. *Suppose $q \in \left(\frac{2}{3}, 1\right)$ and candidates play symmetric reform equilibrium strategies (i.e., separating) when the interest group can bribe. Then the interest group prefers to self-regulate so it cannot bribe if $\pi((1-q)(1-\alpha_1)) - (1-\pi)^2(1-\kappa) \geq 0$, which is satisfied so long as the probability a given candidate is sincere, π, is sufficiently high.*

Proof of Proposition 7. Recall from the proof of Lemma 4 that the group's welfare from self-regulating and inducing access equilibrium candidate behavior is given by,

$$\begin{aligned} W_G(\text{No bribes}|\text{Access}) &= \pi(q(0) + (1-q)(1-\alpha_1)) + (1-\pi)(0), \\ &= \pi((1-q)(1-\alpha_1)). \end{aligned}$$

Suppose instead that the interest group were to choose to keep the ability to bribe. In this environment, when bribery is allowed, and sincere candidates ban access, any time a sincere candidate is running the voter learns who is sincere and corrupt and a sincere politician wins office. The only time in a reform equilibrium with bribery aids the interest group is when two corrupt candidates run against one another since this is the only instance in which the voter will elect a corrupt politician. The group's ex ante expected welfare when bribery is allowed and candidates separate with their access decisions, revealing their types, is given by,

$$\begin{aligned} W_G(\text{Bribes}|\text{Reform}) &= \pi^2(0) + (2(1-\pi)\pi)(0) + (1-\pi)^2(1-\kappa), \\ &= (1-\pi)^2(1-\kappa). \end{aligned}$$

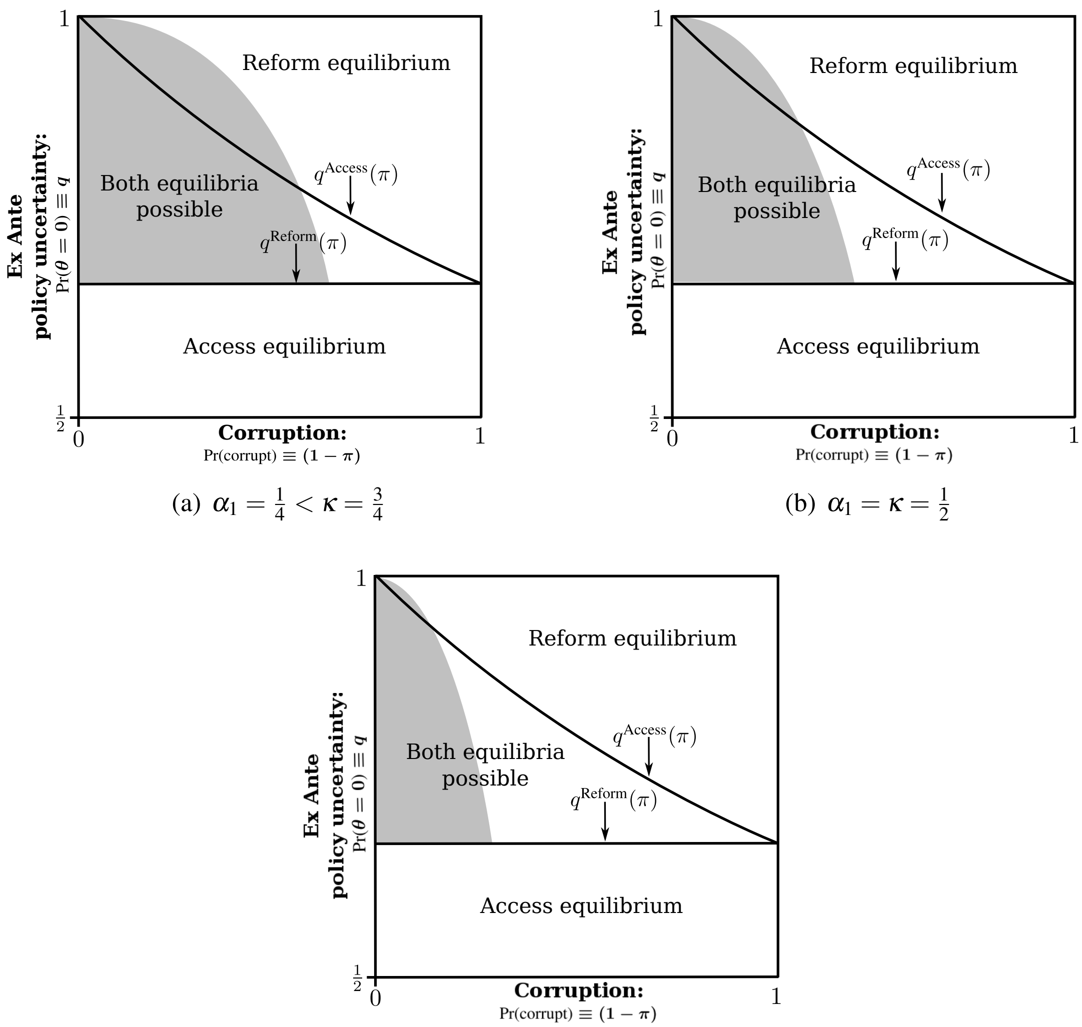
With probability $(\pi^2 + 2(1-\pi)\pi)$ a sincere candidate runs for and wins office, but since that candidate won office by separating and banning access the group cannot lobby and therefore, $x = 0$ always and the group receives zero no matter what. With probability $(1-\pi)^2$ both candidates running are corrupt and therefore the winning politician will be corrupt. In this case the group pays the bribe at cost κ in exchange for implementing $x = 1$, leading to a payoff of $1-\kappa$. Comparing the two welfare expressions in this case yields,

$$W_G(\text{No bribes}|\text{Access}) - W_G(\text{Bribes}|\text{Reform}) = \pi((1-q)(1-\alpha_1)) - (1-\pi)^2(1-\kappa)$$

Thus, an interest group would prefer to self-regulate and "tie its own hands" by ex ante committing to no bribery in a separating reform equilibrium environment so long as,

$$\pi((1-q)(1-\alpha_1)) - (1-\pi)^2(1-\kappa) > 0, \tag{1}$$

which is satisfied for all $q \in \left(\frac{2}{3}, 1\right)$ when $\frac{\alpha_1(1-q)+2k+q-3}{2(\kappa-1)} - \frac{1}{2}\sqrt{\frac{(\alpha_1-1)(q-1)(\alpha_1(q-1)-4\kappa-q+5)}{(\kappa-1)^2}} < \pi < 1$. Therefore, an interest group benefits from self-regulation in a reform equilibrium environment so long as π is sufficiently high. ■



(a) $\alpha_1 = \frac{1}{4} < \kappa = \frac{3}{4}$

(b) $\alpha_1 = \kappa = \frac{1}{2}$

(c) $\alpha_1 = \frac{3}{4} > \kappa = \frac{1}{4}$

Figure 1: Examples of when the interest group prefers no bribery in a separating reform equilibrium environment, conditional on the cost of lobbying relative to the cost of bribery.

Note: The gray shaded area represents the region in which the interest group benefits from bribes being banned given a separating reform equilibrium. Figure 3(a) is an example when lobbying is less costly than bribery. Figure 3(b) is an example when lobbying and bribery are equally costly. Figure 3(c) is an example when lobbying is more costly than bribery.

Figure 1 displays examples of the regions in which the interest group prefers to self-regulate, conditional on the relative costs of lobbying versus bribery. Within the gray shaded region in each example the group would prefer to commit to no bribery given that candidates play separating reform equilibrium strategies. As the figure illustrates, the region in which the group benefits from this self-regulation grows larger as bribery becomes more costly relative to substantive lobbying.

The result in Proposition 7 only directly applies when candidates play a separating reform equilibrium for sure. Much of the region in which equation 1 is satisfied is also the region in which both the reform and access equilibrium are possible. We also know that in a pooling access environment the interest group *always* benefits from being allowed to bribe (from Lemma 3). Thus, committing to no bribery can be beneficial in a separating equilibrium but it is costly in a pooling access equilibrium. So to fully explore when the interest group benefits from committing to no bribery we must take into account both possibilities. There is no *prima facie* reason to expect one equilibrium is more likely to obtain than the other when both are possible so we take an agnostic view and simply assign complementary probabilities to each one to represent the interest group's beliefs about which equilibrium would be played. The following corollary extends the logic in Proposition 7 to this environment by noting that as long as the interest group believes a separating reform equilibrium is sufficiently likely to be played relative to a pooling access equilibrium it will commit to no bribery.

Corollary 2. *Suppose both the reform equilibrium and access equilibrium are possible. Define* $\beta \equiv Pr(\textit{Reform equilibrium})$ *and* $1-\beta \equiv Pr(\textit{Access equilbrium})$*. Then so long as the reform equilibrium is sufficiently likely relative to the access equilibrium the interest group will self-regulate by committing to not bribing corrupt politicians.*

Proof of Corollary 2. Lemma 3 shows that the interest group never wants to self-regulate when candidates play access equilibrium strategies for sure. Proposition 7 shows that there are conditions in which the interest group would prefer to self-regulate and commit to no bribery when candidates play reform equilibrium strategies for sure. Continuity of the interest group's expected utilities with respect to probabilities, derived in the proof of Proposition 7, implies that if the reform equilibrium is sufficiently likely relative to the access equilibrium – i.e., $\frac{\beta}{1-\beta}$ is sufficiently large – then the interest group will still prefer to self-regulate. ■

B Robustness

In this section we discuss existing empirical research that points to candidate commitments to platforms or policies as a reasonable assumption in electoral environments. We also provide two alternative modeling approaches: (1) a model with no politician commitment that leads to similar political behavior, and (2) a model with an alternative specification for state-dependent informational lobbying costs. First, we argue that allowing candidates to commit to a platform provides a better empirical approximation to real elections than would a similarly parsimonious model with no commitment. Second, we explore an alternative model that drops commitment and discuss its relation to our main model in-text. Finally, we show that our key results are robust to an alternative specification of state-dependent lobbying costs: $\alpha_1 < \alpha_0$ rather than the in-text specification requiring that $0 < \alpha_1 < 1 < \alpha_0$. All of these exercises are meant to demonstrate that the key substantive insights we provide in the manuscript are well supported by previous empirical and theoretical literature, and robust to alternative modeling specifications.

B.1 Empirical support for commitment

In a one-shot election with no commitment we would predict that campaign announcements would be meaningless and voters would interpret them as such.[3] However, existing work on campaign promises finds them to be very credible in practice. One line of research on campaign promises shows that Presidents keep or attempt to keep the vast majority of promises they made during the campaign (e.g., Krukones, 1984; Fishel, 1985). Claibourn (2011) updated this research and reached similar conclusions. Sulkin (2009) analyzed campaign promises made in television advertisements by House candidates over three elections and found that "candidates do indeed follow through on the appeals they make in campaigns" (1093). Sulkin (2011) is a book-length follow up on the above article that expands on the argument by including additional congressional data and case

[3]There are notable exceptions to this rule which do not apply in this context. For instance, Schnakenberg (2016) constructs partially informative directional cheap talk equilibria in one-shot elections, but these rely on elements of the game which do not exist here (uncertainty about the voter's type, indifference of the voter between some politician types, etc.).

studies of defense and the environment. Her conclusion is that "there is no evidence that legislators regularly claim to be in favor of particular policies but then work against them" (196). Pétry and Collette (2009) perform a meta-analysis of different studies of whether parties in Europe and North America keep their promises and find that they do so over two-thirds of the time on average, concluding that "[c]ontrary to popular belief, political parties are reliable promise keepers." Tomz and Van Houweling (2012*a*,*b*) demonstrate evidence of one mechanism for credibility of campaign promises and show that voters punish politicians who flip-flop or break promises.

Though a variety of mechanisms may contribute to the high overall levels of promise-keeping among politicians and a model without commitment may allow us to explore these mechanisms further, the most realistic baseline assumption, based on a large body of empirical research, is that politicians treat their promises as binding rather than that they feel free to ignore them as soon as they take office. To be clear, none of these papers focus specifically on the choice of whether or not to grant access to lobbyists. However, since this paper is the first (to our knowledge) on the topic of campaign commitments about dispositions toward lobbyists, we think a good starting point is to treat it like every other campaign promise.

B.2 Alternative model with no commitment

The presence or absence of commitment is one of the traditional lines along which models of elections are categorized.[4] The oldest tradition in political economy is to analyze models with strong commitment assumptions and the most well-known models fall into this category (Downs, 1957; Hotelling, 1929; Calvert, 1985; Wittman, 1983). Our model differs from these models in that it does *not* have candidates commit to a policy but instead to a process by which the policy is chosen. A major break from models with commitment was the citizen-candidate models starting with Osborne and Slivinski (1996); Besley and Coate (1997). The existing literature provides some clues about which aspects of the model should generalize to a setting without commitment.

Two of our substantive results are of interest here. First, the threat of corruption may lead non-corrupt politicians to signal to voters that they are not corrupt by denying access to lobbyists even

[4] Ashworth (2012) and Dewan and Shepsle (2011) provide thorough overviews of electoral models.

though they would be more effective policymakers if they did not do so. Second, this incentive may be observed *ex ante* during the campaign phase, where politicians have an incentive to make credible promises. We focus on demonstrating that the first insight extends easily to a model with no commitment by politicians, and therefore no campaign stage. We do not try to extend the second insight to a setting with no commitment. It is well-known in the literature that, in most simple settings, a lack of commitment renders campaign statements useless. For instance, in the literature on citizen-candidate models we find frequent statements like the following from Banks and Duggan (2008):

> As in the citizen-candidate literature, we view campaign promises as cheap talk (and therefore omit them from the model), and elected representatives choose policy unconstrained by past commitments.

Of course, in real elections, politicians frequently make campaign promises and voters act as if the politicians mean what they say (as demonstrated by the empirical literature discussed above). Thus, on this score, this is a point in favor of commitment-based models with respect to empirical realism. In a more complex dynamic model, the appearance of commitment may arise from reputational effects in equilibrium (Aragonés, Palfrey and Postlewaite, 2007). We do not explore this possibility but note that the original model presented in text could represent a reduced version of such a dynamic model.

With this in mind, we now turn to the *No Commitment Game*. There are two periods, 1 and 2. Period one proceeds as follows. There is an Incumbent in power in the first period. Nature draws the Incumbent's type $\tau_I \in \{S, C\}$ at random, with $\Pr[\tau_I = S] = \pi$. Nature also draws the state of the world $\theta_1 \in \{0, 1\}$, with $\Pr[\theta_1 = 0] = q > \frac{1}{2}$. There are no campaign announcements: such announcements are non-credible cheap talk in a model with no commitment and are therefore omitted. Instead, the Incumbent in power chooses whether to grant or deny access while in office in period one. The voter observes this choice. The lobbyist's strategy should not change in any significant way. Thus, we simplify the model by omitting the lobbyist and plugging in the outcomes from the lobbying stage of the static game. Thus, if a sincere type takes office and grants access, policy is set equal to $x = \theta$ in every state. If a sincere type takes office and does not grant access the policy is set equal to $x = 0$. If a corrupt type takes office and grants access the policy is set equal to $x = 1$ and the corrupt politician gets a payoff of one. Finally, if a corrupt type takes off and does not grant access (which will never happen on the path of play), policy is set equal to $x = 0$ and the corrupt politician gets a payoff of zero. This concludes the first period.

Period two proceeds as follows. The Voter observes the access choice of the politician and updates about the Incumbent's type.[5] The Voter then decides whether to retain the Incumbent or replace her with the Challenger. Next, Nature draws a new state of the world θ_2 from the same distribution as θ_1. If the Voter chose the Challenger, Nature also draws the challenger's type τ_C independently from the same distribution as τ_I. That is, $\Pr[\tau_C = S] = \pi$.

Politician stage game payoffs, $i \in \{I, C\}$ and $t \in \{1, 2\}$, are the same as in the main text:

$$u_{it}(\theta, x, b) = \begin{cases} -|x-\theta| & \text{if } \tau_i = S \\ bx - (1-b)x & \text{if } \tau_i = C. \end{cases} \tag{2}$$

Thus, politicians' overall payoffs are given by,

$$u_{i1}(\theta, x, b, \delta) + \delta u_{i2}(\theta, x, b, \delta),$$

where $\delta \in (0, 1)$ is a common discount factor so that as $\delta \to 1$ politicians equally weight payoffs across periods and as $\delta \to 0$ politicians fully weight payoffs in $t = 1$. It is immediately clear that in the terminal stage, $t = 2$, any politician, regardless of type, will grant access. For sincere types they are always better off granting access and being able to match policy to the state. For corrupt types they are always better off granting access and being bribed. In both cases this follows from

[5] We allow the Voter to observe only the access decision to make the updating as simple as possible. Allowing the Voter to also observe policy would make no difference with respect to whether or not there exists a separating equilibrium of the type we will describe, since in pure strategies the Voter will update beliefs to either $\Pr[\tau_I = S] = 1$ or $\Pr[\tau_I = S] = 0$ along the path of play. That is, any information set at which the politician reneges on her campaign promise is off the path of play and therefore not pinned down by Bayes Rule under Perfect Bayesian Equilibrium. Allowing the Voter to also observe θ_1 may complicate matters, but realistically it is difficult for a Voter in a realistic setting to attribute her payoff to the effects of any particular policy.

the fact that there are no future considerations to take into account (i.e., the voter can not discipline behavior in the terminal period). This is standard in finitely repeated games of this sort.

We now turn to analyzing the conditions under which we can support a separating equilibrium (qualitatively analogous to the separating reform equilibrium in the main text). In a separating equilibrium to this *No Commitment Game* sincere incumbents always ban access in the first period and corrupt incumbents always grant access in the first period. This allows the Voter to learn which incumbents are sincere and corrupt. In turn, the Voter's best response is to remove any incumbent who chose to grant access in the first period in favor of a random challenger in period two and retain any incumbent who chose to ban access in the first period. This follows from the fact that all politicians grant access period two, which implies that the Voter always wants to screen out corrupt politicians that will invariably set $x_2 = 1$ in order to retain only sincere politicians who will match policy to the state. Given this voter strategy sincere politicians must weigh their payoffs for granting access today, thereby matching $x_1 = \theta_1$, and having a challenger choose policy in period two against banning access today, which means a potential mismatch of policy and state in period one, but remaining in office in period two, which will lead to matching $x_2 = \theta_2$. Corrupt politicians, on the other hand, decide between granting access today, which means they will be removed from office in period two, and mimicking sincere incumbents by banning access today, which allows them to grant access tomorrow. We first show that corrupt politicians always prefer to grant access in period one and accept being removed from office in period two, rather than ban access in period one and grant access in period two.

Consider a corrupt Incumbent's payoff from playing his equilibrium strategy of granting access in period one (knowing he will be removed in period two),

$$
\begin{aligned}
EU_I(\text{grant in period } 1 | \tau_I = C) &= bx_1 + (1-b)x_1 + \delta\left[bx_2 + (1-b)x_2\right], \\
&= (1)(1) + (1-1)(1) + \delta\left[0\right], \\
&= 1.
\end{aligned}
$$

If the corrupt Incumbent were to instead ban access in period one, which leads to him being retained and granting access in period two, his payoff would be,

$$\begin{aligned} EU_I(\text{ban in period } 1|\tau_I = C) &= (0)0 + (1-0)0 + \delta[1], \\ &= \delta. \end{aligned}$$

Thus, corrupt incumbents strictly prefer to grant access in period one for all $\delta \in (0,1)$: they have a dominant strategy of granting access in the first period even though they will be removed in period two given the Voter's strategy.

Now consider a sincere Incumbent's expected payoff for playing his equilibrium strategy of banning access in period one, which leads to him being retained, and granting access in period two,

$$\begin{aligned} EU_I(\text{ban in period } 1|\tau_I = S) &= -|x_1 - \theta_1| - \delta[|x_2 - \theta_2], \\ &= -(q|0-0| + (1-q)|0-1|) - \delta[|\theta_2 - \theta_2|], \\ &= -(1-q) \end{aligned}$$

If instead the sincere Incumbent were to grant access in period one, he would be removed from office in period two in favor of a random Challenger, leading to the following expected payoff,

$$\begin{aligned} EU_I(\text{grant in period } 1|\tau_I = S) &= -|\theta_1 - \theta_1| - \delta[\pi|\theta_2 - \theta_2| + (1-\pi)(q|1-0| + (1-q)|1-1|)], \\ &= -\delta(1-\pi)q. \end{aligned}$$

In order for banning access in period one to be incentive compatible for sincere incumbents the following inequality must be satisfied:

$$
\begin{aligned}
-(1-q) &\geq -\delta(1-\pi)q, \\
q &\geq \frac{1}{1+\delta(1-\pi)}.
\end{aligned}
$$

Thus, corrupt incumbents always grant access in period one and sincere incumbents ban access in period one, for all $\delta \in (0,1)$ and $\pi \in (0,1)$, if and only if $q \in \left[\frac{1}{1+\delta(1-\pi)}, 1\right]$. So long as this condition on q (as a function of π and δ) holds then we have a qualitatively analogous separating equilibrium in the *No Commitment Game*.

Proposition 8. *For all $\delta \in (0,1)$ and $\pi \in (0,1)$, if $q \in \left[\frac{1}{1+\delta(1-\pi)}, 1\right]$ then there is a separating equilibrium to the* No Commitment Game *characterized by the following strategies:*

- *Sincere incumbents ban access in period one and grant access in period two,*
- *Corrupt incumbents grant access in period one (and, off-path, grant access in period two),*
- *The voter retains incumbents in period two if and only if they banned access in period one.*

Proof of Proposition 8. Follows directly from the analysis above. ■

Notice that $\lim_{\delta\to 1}\left(\frac{1}{1+\delta(1-\pi)}\right) = \frac{1}{2-\pi}$ and $\lim_{\delta\to 0}\left(\frac{1}{1+\delta(1-\pi)}\right) = 1$, implying that as δ decreases, and the Incumbent becomes more myopic, it is more difficult to support separating equilibrium. This constraint only binds sincere incumbents. Corrupt incumbents strictly prefer to grant access in the first period so long as there is any discounting of the future. This is intuitive: the more myopic the sincere incumbent is the stronger are his incentives to grant access in period one to match policy to the state rather than ban access to ensure he remains in office in period two. Conversely, the more forward-looking the sincere incumbent is the stronger are his incentives to sacrifice expected policy payoffs by banning access in period one in order to ensure he remains in office in period two, in which case he will grant access and match policy to the state.

Overall, the foregoing analysis shows that the key result from the main model of this article – the threat of corruption provides incentives for non-corrupt politicians to ban access in order to signal their sincerity – generalizes to this alternative model with no commitment. This implies that the substantive results presented in the main text are not wholly dependent on the ability of politicians to credibly commit to campaign promises. We have shown that under this alternative specification we can support the same substantive behavior from sincere and corrupt politicians (i.e., separating) without that commitment.

B.3 Alternative state-dependent lobbying costs

In this section we provide an alternative signaling technology for informational lobbying and show that we can still support interest group separation with informational lobbying decisions. Specifically, we hold every aspect of the in-text model fixed *except* that rather than $m \in \{0,1\}$ we now make informational lobbying continuous so that $m \geq 0$ and rather than $\alpha_0 > 1 > \alpha_1 > 0$ we simply assume that $\alpha_0 > \alpha_1 > 0$.

We focus our analysis only on the interest group-winning politician interactions characterized in Proposition 1 in the main body of the paper as this is the only stage where behavior is affected by these alterations to m and the nature of the ordering of α_0 and α_1. First, note that since corrupt politicians only respond to b and not m their behavior as described in Proposition 1 remains unchanged. What we want to show then is that informational lobbying of sincere politicians as described in-text holds under this alternative model specification: the interest group reveals the true state θ to the politician by separating with its choice of m. This implies that the good interest group ($\theta = 1$) chooses a level of informational lobbying, which we denote with m^*, that deters the bad interest group ($\theta = 0$) from informational lobbying, which requires that this type, at best, is made indifferent between choosing m^* and $m = 0$. The following result formally establishes that this separating equilibrium exists under this alternative specification.

Proposition 9. *Assume the model described in the main text, but let* $m \geq 0$ *and* $\alpha_0 > \alpha_1 > 0$ *(instead of* $m \in \{0,1\}$ *and* $\alpha_0 > 1 > \alpha_1 > 0$*). Further, suppose a sincere politician wins office. There exists a feasible* $m^* > 0$ *such that the interest group separates with its informational lobbying choices and reveals* θ *to the politician.*

Proof of Proposition 9. Suppose that the interest group separates with its informational lobbying choice, m. Then when $\theta = 1$ the politician implements $x^*(m) = 1$ and when $\theta = 0$ the politician implements $x^*(m) = 0$ by the same arguments in the proof of Proposition 1. To show that this is a best response for the interest group consider group utility in this case:

$$u_G(m, b = 0|\tau_i = S, x^*, \theta) = \begin{cases} x^* - \alpha_1 m & \text{if } \theta = 1, \\ x^* - \alpha_0 m & \text{if } \theta = 0. \end{cases}$$

Interest group separation requires that the good type of interest group ($\theta = 1$) strictly benefits from the choice of m while the bad type of interest group ($\theta = 0$) is, at best, indifferent. Denote the level of informational lobbying that satisfies this requirement as m^*. Leveraging $x^* = 1$ when $\theta = 1$ and $x^* = 0$ when $\theta = 0$ in a separating equilibrium and plugging these into the payoff expressions above, these requirements further imply that the following set of inequalities must hold simultaneously,

$$\begin{aligned} 1 &> \alpha_1 m^*, \\ 1 &\leq \alpha_0 m^*. \end{aligned}$$

The first inequality ensures that the good type benefits from the choice of m^* and the second ensures that the bad type is at best indifferent when mimicking the good type (which induces $x^* = 1$). These two inequalities are satisfied simultaneously when,

$$m^* \in \left[\frac{1}{\alpha_0}, \frac{1}{\alpha_1}\right).$$

Finally, note that $\frac{1}{\alpha_1} > \frac{1}{\alpha_0}$ since $\alpha_0 > \alpha_1$. This implies that there always exists a m^* in that range that the good type of interest group can select that deters the bad type from mimicking. Thus, when $m \geq 0$ and $\alpha_0 > \alpha_1 (> 0)$ the interest group will separate with its informational lobbying decisions following a sincere candidate taking office, as was to be shown. ■

References

Aragonés, Enriqueta, Thomas Palfrey and Andrew Postlewaite. 2007. "Political Reputations and Campaign Promises." *Journal of the European Economic Association* 5(4):846–884.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15:183–201.

Banks, Jeffrey S. and John Duggan. 2008. "A Dynamic Model of Democratic Elections in Multidimensional Policy Spaces." *Quarterly Journal of Political Science* 3(3):269–299.

Besley, Timothy and Stephen Coate. 1997. "An Economic Model of Representative Democracy." *Quarterly Journal of Economics* 108:85–114.

Calvert, Randall L. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science* 29(1):69–95.

Claibourn, Michele P. 2011. *Presidential Campaigns and Presidential Accountability*. Urbana, IL: University of Illinois Press.

Dewan, Torun and Kenneth A. Shepsle. 2011. "Political Economy Models of Elections." *Annual Review of Political Science* 14:311–330.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York, NY: Harper.

Fishel, Jeff. 1985. *Presidents and Promises: From Campaign Pledge to Presidential Performance*. Washington, DC: Congressional Quarterly Press.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal* 39(152):41–57.

Krukones, Michael G. 1984. *Promises and Performance: Presidential Campaigns as Policy Predictors*. Lanham, MD: University Press of America.

Osborne, Martin and Al Slivinski. 1996. "A Model of Political Competition with Citizen Candidates." *Quarterly Journal of Economics* 111:65–96.

Pétry, Francois and Benoit Collette. 2009. Measuring How Political Parties Keep Their Promises: A Positive Perspective from Political Science. In *Do They Walk Like They Talk? Speech and Action in Policy Processes*, ed. Louis M. Imbeau. New York, NY: Springer.

Schnakenberg, Keith E. 2016. "Directional Cheap Talk in Electoral Campaigns." *Journal of Politics* 78(2):527–541.

Sulkin, Tracy. 2009. "Campaign Appeals and Legilsative Action." *Journal of Politics* 71(3):1093–1108.

Sulkin, Tracy. 2011. *The Legislative Legacy of Congressional Campaigns*. New York, NY: Cambridge University Press.

Tomz, Michael and Robert P. Van Houweling. 2012*a*. "Candidate Repositioning." *Unpublished Manuscript. Stanford University* .
URL: *https://web.stanford.edu/~tomz/working/TomzVanHouweling-Repositioning-2012-10-24.pdf*

Tomz, Michael and Robert P. Van Houweling. 2012*b*. "Political Pledges as Credible Commitments." *Unpublished Manuscript. Stanford University* .
URL: *https://web.stanford.edu/~tomz/working/TomzVanHouweling-Pledges-2012-03-27.pdf*

Wittman, Donald. 1983. "Candidate Motivation: A Synthesis of Alternative Theories." *The American Political Science Review* 77(1):142–157.