

SUPPLEMENTARY APPENDIX

B1. Unique Equilibrium with Asymmetric Groups

Before proving the results in the paper, we first establish that a unique political equilibrium exists even when groups are asymmetric if k is uniform on $[0, 1]$. For group G , we denote by w_G the importance of the election for members of group G , by F_G the distribution of their voting costs, and by λ_G the strength of social incentives. Let P_G denote the Pragmatist Best Response for group G , and for a cost c_G , let $T_G(c) = qF_G(c) + (1 - q)F_G(P_G(c))$ denote the total turnout from group G ; it follows that T_G is increasing in c . The first-order conditions are

$$(B1) \quad \begin{aligned} \frac{T_2(c_2)}{(T_1(c_1))^2} w_1 h \left(\frac{T_2(c_2)}{T_1(c_1)} \right) &= \frac{c_1}{2}, \\ \frac{1}{T_1(c_1)} w_2 h \left(\frac{T_2(c_2)}{T_1(c_1)} \right) &= \frac{c_2}{2}. \end{aligned}$$

Therefore,

$$(B2) \quad \frac{c_1 T_1(c_1)}{w_1} = \frac{c_2 T_2(c_2)}{w_2}.$$

For the claims that follow, we assume that k is uniformly distributed on $[0, 1]$; the two claims together imply existence and uniqueness of political equilibrium.

CLAIM 1: (c_1, c_2) that satisfy the first-order conditions in Equation B1 are maxima.

PROOF:

As in the proof of Theorem 1, it suffices to establish that $h' \left(\frac{T_2(c_2)}{T_1(c_1)} \right) < 0$, which is necessarily true when $\alpha = 1$ since $h'(x) = -\frac{2}{(x+1)^3} < 0$ for all x .

CLAIM 2: *There is a unique solution to Equation B1.*

PROOF:

Our argument adapts the proof of Fact 1 on p. 22-23 of Feddersen and Sandroni (2006c) to our environment. Suppose that there were two solutions c_1, c_2 and c'_1, c'_2 . It follows from Equation B2 that if $c'_1 > (=, <) c_1$, then $c'_2 > (=, <) c_2$, and so WLOG, we assume that $c'_1 > c_1$. Let $\tau = \frac{T_2(c_2)}{T_1(c_1)}$ and $\tau' = \frac{T_2(c'_2)}{T_1(c'_1)}$. It follows from Equation B2 that

$$\text{sgn}(\tau' - \tau) = \text{sgn} \left(\frac{c'_1}{c'_2} - \frac{c_1}{c_2} \right).$$

We study two complementary cases below and show how each yields a contradiction.

- 1) $\frac{c'_1}{c'_2} \geq \frac{c_1}{c_2}$: Re-writing the second equation in Equation B1 yields that

$$c'_2 T_1(c'_1) = 2w_2 h(\tau') \leq 2w_2 h(\tau) = c_2 T_1(c_1),$$

in which the inequality follows h being a decreasing function if $\alpha = 1$ and $\tau' \geq \tau$. Since T_1 is a strictly increasing function, the above equation contradicts $(c'_1, c'_2) \gg (c_1, c_2)$.

- 2) $\frac{c'_1}{c'_2} < \frac{c_1}{c_2}$: Analogous to h , consider the density function of $\frac{1-k}{k}$ denoted by \tilde{h} . It follows that the first equation of Equation B1 can be re-written as

$$(B3) \quad \frac{1}{T_2(c_2)} w_1 \tilde{h}\left(\frac{1}{\tau}\right) = \frac{c_1}{2}.$$

Since \tilde{h} is decreasing in its argument, and $\tau' < \tau$, it follows that

$$c'_1 T_2(c'_2) = 2w_1 h\left(\frac{1}{\tau'}\right) < 2w_1 h\left(\frac{1}{\tau}\right) = c_1 T_2(c_2),$$

which contradicts $(c'_1, c'_2) \gg (c_1, c_2)$.

B2. Proofs for Section II

PROOF OF THEOREM 2:

We begin by establishing that the Pragmatist Best Response is unique. For $P(c^*)$ as defined in Definition 1, and for arbitrary c^* , consider

$$(1-s)\zeta(1, P(c^*), c^*) - \zeta(0, P(c^*), c^*).$$

The term above is continuously decreasing in s , is strictly positive when $s = 0$, and strictly negative when $s = 1$. Let $\bar{s}(c^*)$ be the unique value of s such that the term above is 0.

First, suppose that $s \geq \bar{s}(c^*)$: we claim that the unique Pragmatist Best Response in this setting is identical to $P(c^*)$ (of Definition 1) and with $\mu_G = 0$. Notice that the Truthful Abstainer's Image is at least as large as the Lying Abstainer's Image, and therefore, no pragmatist has an incentive to deviate. Suppose that there was another pragmatist response in which $\mu_G > 0$ and the pragmatist's cost cutoff is \tilde{c} . To satisfy (4) of Definition 4, it follows that $F(\tilde{c}) + \mu_G \leq F(P(c^*))$.

Because a pragmatist with cost \tilde{c} is indifferent between voting and not,

$$\begin{aligned}\tilde{c} &= \lambda \left(\frac{s \frac{qF(c^*)}{qF(c^*) + (1-q)F(\tilde{c})} + (1-s) \frac{qF(c^*)}{qF(c^*) + (1-q)(F(\tilde{c}) + \mu_G)}}{q(1-F(c^*))} \right) \\ &\geq \lambda (\zeta(1, P(c^*), c^*) - \zeta(0, P(c^*), c^*)) \\ &= P(c^*),\end{aligned}$$

which is a contradiction.

Now, suppose that $s < \bar{s}(c^*)$: for all $s \leq \bar{s}(c^*)$, find the unique cost, $c(s, c^*)$, that makes a voter indifferent between voting and obtaining the Voter's Image, and abstention-lying obtaining the Lying Abstainer's Image when ethical citizens use a cutoff of c^* . It follows that

$$c(s, c^*) = \lambda s \frac{qF(c^*)}{qF(c^*) + (1-q)F(c(s, c^*))}.$$

For the equality to hold, $c(s, c^*)$ is increasing in s , and by construction, $c(\bar{s}(c^*), c^*) = P(c^*)$. Therefore, for $s < \bar{s}(c^*)$, $c(s, c^*) < P(c^*)$. Setting $\tilde{c}_G = c(s, c^*)$ consider the expression given by Lying Abstainer's Image – Truthful Abstainer's Image: it is continuously decreasing in μ_G and strictly negative at $\mu_G = 1 - F(\tilde{c}_G)$. Moreover, at $\mu_G = 0$, the term is

$$\begin{aligned}&(1-s)\zeta(1, c(s, c^*), c^*) - \zeta(0, c(s, c^*), c^*) \\ &> (1 - \bar{s}(c^*))\zeta(1, c(s, c^*), c^*) - \zeta(0, c(s, c^*), c^*) \\ &> (1 - \bar{s}(c^*))\zeta(1, P(c^*), c^*) - \zeta(0, P(c^*), c^*) \\ &= 0,\end{aligned}$$

in which the first inequality follows from $s < \bar{s}(c^*)$, the second inequality follows from $c(s, c^*) < P(c^*)$, and the equality is by construction. Therefore, there exists a unique μ_G that equates Lying Abstainer's Image and Truthful Abstainer's Image.

Thus, we have constructed the unique Pragmatist Best Response in the model that permits lying. Based on the above discussion, we extend $c(s, c^*)$ to the entire domain $[0, 1] \times [0, \infty)$; since signaling incentives are increasing in c^* , it follows that $c(s, c^*)$ is increasing in c^* . The first-order condition generated by the Consistent Ethical Rule is similar to that in Theorem 1, and yields an analogue to Equation 4:

$$c_1^* (qF(c_1^*) + (1-q)F(c(s, c_1^*))) = c_2^* (qF(c_2^*) + (1-q)F(c(s, c_2^*))).$$

Since $c(s, c^*)$ is increasing in c^* , it follows that $c_1^* = c_2^*$, from which the characterization Theorem 2 follows.

Consider the ethical cutoff from the unique political equilibrium in Theorem 1 and denote this by C . Let $\tilde{s} = \bar{s}(C)$. It follows that if $s \geq \tilde{s}$, the unique equilibrium corresponds to Theorem 1. On the other hand, if $s < \tilde{s}$, then the unique equilibrium involves $\mu_G > 0$, and a pragmatist participation less than that in the setting without lying $c(s, c^*) < P(C)$. As s increases, $c(s, c^*)$ increases and so it follows that c_G^* decreases for each group G while overall turnout increases.

PROOF OF THEOREM 3:

We first establish the existence and uniqueness of a pragmatic best response. Holding fixed an ethical participation rate μ^* , notice that $\zeta(1, c, \mu^c, \mu, \mu^*)$ is strictly decreasing in μ^c and $\hat{\mu}$ and $\zeta(0, c, \mu, \mu^*)$ is strictly increasing in $\hat{\mu}$. For each c , and for each $\mu \in [0, 1]$, let $h^c(\mu, \mu^*)$ define the participation rate among pragmatists with cost c when the participation rates among all pragmatists and ethical citizens in group G are μ and μ^* respectively. Formally, for $c \leq c^*$, $h^c = 1$ if $c \leq \lambda(\zeta(1, c, 1, \mu, \mu^*) - \zeta(0, c, \mu, \mu^*))$, $h^c = 0$ if $c > \lambda(\zeta(1, c, 0, \mu, \mu^*) - \zeta(0, c, \mu, \mu^*))$, and otherwise,

$$h^c(\mu, \mu^*) = \left[(c/\lambda + \zeta(0, c, \mu, \mu^*) - \zeta(0, c, 1 - \mu, 1 - \mu^*))^{-1} pq - q \right] / (1 - q).$$

In essence, h^c is 1 (resp. 0) if a citizen with cost c is better off voting (resp. abstaining) even when it is known that all citizens with that cost vote (resp. abstain). Otherwise, h^c finds the randomization probability that makes a voter with cost c indifferent. It follows that h^c is weakly decreasing and continuous in μ .

For $\mu \in [0, 1]$, let $h(\mu, \mu^*) = \int_0^{\bar{c}} h^c(\mu, \mu^*) dF$. Since $h(\mu, \mu^*) \in [0, 1]$, and h is continuous and weakly decreasing in μ , it follows that there exists a unique μ that satisfies $\mu = h(\mu, \mu^*)$. This is the unique Pragmatist Best Response. The remainder of the argument for existence and uniqueness follows Theorem 1.

B3. Proofs for Section III

PROOF OF THEOREM 4:

We first establish that $\frac{3}{2} - x$ is not an equilibrium platform because it is defeated by each of the other platforms with probability greater than 1/2:

$$\frac{1}{\kappa \kappa_x} = \frac{w_1\left(\frac{3}{2}, \frac{3}{2} - x\right)}{w_2\left(\frac{3}{2}, \frac{3}{2} - x\right)} < \frac{w_1\left(\frac{3}{2} + x, \frac{3}{2} - x\right)}{w_2\left(\frac{3}{2} + x, \frac{3}{2} - x\right)} = \frac{1}{\kappa} \leq 1.$$

Therefore, by Property 5, $\frac{3}{2} - x$ is defeated by each of the other platform with probability strictly greater than $\frac{1}{2}$ if $\kappa > 1$, and by Property 6 if $\kappa = 1$ and $\lambda_2 > \lambda_1$.

When responsiveness is asymmetric, $\frac{3}{2}$ defeats (resp. is defeated by) $\frac{3}{2} + x$ if $\kappa <$ (resp. $>$) κ_x with probability exceeding $\frac{1}{2}$, yielding the unique equilibrium

prediction. When social incentives are asymmetric, notice that if $\kappa_x = 1$ then by Property 6 it follows that $\lambda_2 > \lambda_1$ implies that $\frac{3}{2} + x$ defeats $\frac{3}{2}$ with probability exceeding $\frac{1}{2}$. Therefore, the desired result follows by continuity.

PROOF OF THEOREM 5:

For proposals (p_1, p_2) , let $\Pi(p_1, p_2)$ denote the probability with which candidate 1 wins. Candidate G's payoff therefore is:

$$\Pi(p_1, p_2)v(|p_1 - G|) + (1 - \Pi(p_1, p_2))v(|p_2 - G|).$$

Since $\Pi(p_1, p_2) \in (0, 1)$ for every (p_1, p_2) , it follows that in every pure strategy equilibrium, candidates run on different platforms. Moreover, notice that if $(p_1, p_2) = (\frac{3}{2} - x, \frac{3}{2} + x)$, then Candidate 2 wins with probability of at least $\frac{1}{2}$, and therefore, $(\frac{3}{2}, \frac{3}{2} + x)$ is not an equilibrium. When responsiveness is asymmetric, then the unique equilibrium platform is $(\frac{3}{2}, \frac{3}{2} + x)$ if $v_x = 1$ and $\kappa > \kappa_x$, and therefore the result follows by continuity. When social incentives are asymmetric, then the unique equilibrium platform is $(\frac{3}{2}, \frac{3}{2} + x)$ if $v_x = 1$, $\kappa_x = 1$ and $\lambda_2 > \lambda_1$, and therefore the result follows by continuity.

B4. Naïve Ethics

The naïve ethical determines behavior as if all citizens are ethical, and therefore (incorrectly) deems the social cost of voting to be

$$\phi^N(c_1, c_2) = E[k] \int_0^{c_1} c dF + (1 - E[k]) \int_0^{c_2} c dF.$$

The aggregate welfare as perceived by group 1 (that for group 2 is analogous) is

$$(B4) \quad V_1^N(c_1, c_2) = w \left(1 - H \left(\frac{F(c_2)}{F(c_1)} \right) \right) - \phi(c_1, c_2).$$

DEFINITION 5: A profile (c_1^N, c_2^N) is a **Naïve Ethical Rule** if for every group G ,

$$V_G^N(c_G^N, c_{-G}^N) \geq V_G^N(c, c_{-G}^N) \text{ for all } c > 0.$$

In contrast to Definition 2, Naïve Ethical Rules are not “best-responses” to the behavior of pragmatists but to the ethical rule of the opposing group. A naïve political equilibrium then is simply a profile of thresholds $\{(c_G^N, \hat{c}_G)_{G=1,2}\}$ such that (c_1^N, c_2^N) are Naïve Ethical Rules, and $\hat{c}_G = P(c_G^N)$. Analogous to Theorem 1, it is straightforward to show that the following holds.

THEOREM 6: *There is a unique naïve political equilibrium: for every group G , c_G^N solves*

$$(B5) \quad cF(c) = \frac{2^{1-2\alpha}w}{B(\alpha, \alpha)}.$$

By comparison to the political equilibrium described in Section I.C, the naïve political equilibrium necessarily has a lower participation rate among ethical citizens and pragmatists because ethical citizens do not compensate for the lower participation of pragmatists. It is straightforward to show that Properties 1 and 3 apply; because naïve ethical citizens do not respond to pragmatists, only those parts of Properties 2 and 6 regarding average participation continue to hold.

B5. Continuum of Types

Given an ethical rule, let $R_G(a_i, c_i)$ be a binary indicator that is 1 if and only if the action a_i is ethical when the cost of voting is c_i . Each citizen privately knows how much she values following the ethical rule; her *ethical coefficient* D_i scales her private gain from behaving ethically and is drawn from the interval $[0, \infty)$ with a smooth cdf F^E , independent of her voting cost; as before, the individual's voting cost c_i is drawn from $[0, \infty)$. Within this setting, a citizen i is ethical if $D_i > c_i$ since she is willing to vote or abstain as required by the ethical rule.¹

When an agent i belongs to group G , her payoff from taking action a_i is

$$(B6) \quad -c_i a_i + D_i R_G(a_i, c_i) + \lambda \Pr(D_i \geq c_i | a_i).$$

This formulation encapsulates intrinsic motivation towards ethics, as captured by D_i , and the extrinsic motivation to appear to be a member of this ethical group, as captured by λ .

Figure B1 describes voting behavior in this context. One can partition the citizens types (D_i, c_i) into three categories: those who abstain, those who vote because of social incentives, and ethical citizens who vote absent social incentives. The vertical line at c_G^* represents the cost cutoff from the ethical rule, and the 45 degree line through the origin separates the citizens in the latter two categories. The gap between the two 45 degree lines illustrates the extrinsic motives for voting, and all other types abstain. As in the binary types model, no citizen with costs above the ethical cut-off chooses to vote.

THEOREM 7: *There exists a unique political equilibrium.*

PROOF:

¹Were voting costs to have an upper-bound, \bar{c} , a simpler definition of an ethical citizen would be one for whom $D_i > \bar{c}$. To maintain consistency with the rest of the paper, we model costs as being unbounded.

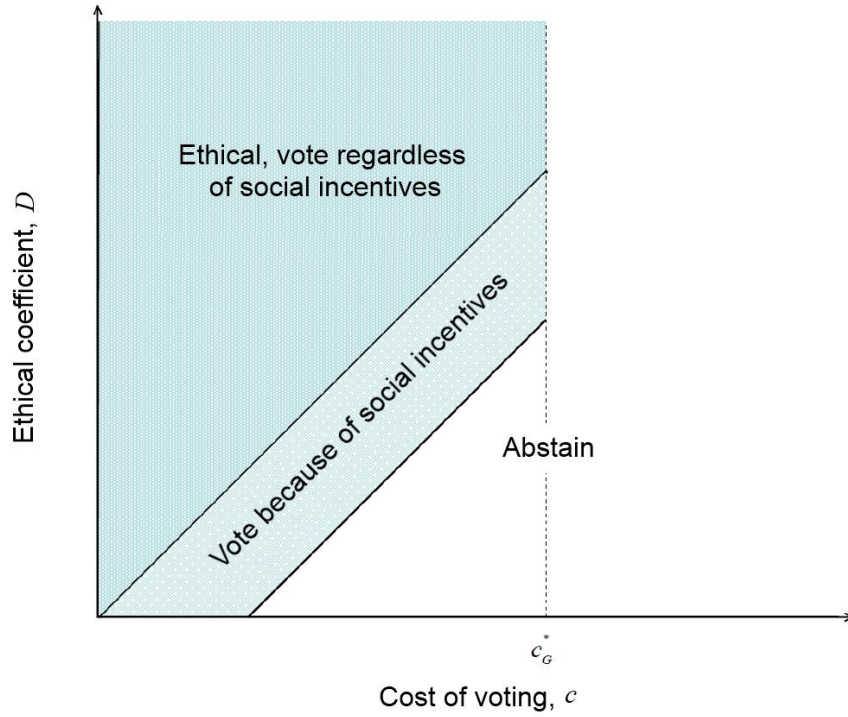


FIGURE B1. VOTING BEHAVIOR WITH A CONTINUUM OF TYPES

First, we show that there exists a unique Pragmatic Best Response given any ethical rule c_G^* . Fix a given c_G^* , and for a fixed S , consider the sets of types:

$$\Gamma_S = \left\{ (D_i, c_i) : D_i (R_G(1, c_i) - R_G(0, c_i)) + S \geq c_i \right\},$$

$$\tilde{\Gamma}_S = \mathfrak{R}_+^2 \setminus \Gamma_S.$$

The sets above consider those types that vote (resp. abstain) when the extrinsic incentive corresponds to S . Let

$$\iota(S) = \lambda \left(\Pr \left[D_i \geq c_i | \Gamma_S \right] - \Pr \left[D_i \geq c_i | \tilde{\Gamma}_S \right] \right)$$

denote the payoff gap induced from social esteem between voters and abstainers when it is believed that only types in Γ_S vote. A Pragmatist Best Response satisfies $\iota(S) = S$.

Towards showing that a Pragmatist Best Response exists, observe that $\iota(0) > 0$: Γ_0 comprises types above the 45 degree line in Figure B1, and therefore, $\Pr[D_i \geq c_i | \Gamma_0] = 1$. In contrast, $\tilde{\Gamma}_0$ comprises types below the 45 for costs below c_G^* , and therefore, $\Pr[D_i \geq c_i | \tilde{\Gamma}_0] < 1$. Now, observe that $\iota(c_G^*) \leq 0$: $\Gamma_{c_G^*}$ comprises all types for which $c_i \leq c_G^*$, and $\tilde{\Gamma}_{c_G^*}$ comprises all types for which $c_i > c_G^*$. Because $\Pr[D_i \geq c_i | c_i]$ is decreasing in c_i , it follows that $\iota(c_G^*) \leq 0$. Finally, we note that $\iota(S)$ is strictly decreasing in S . Let $S^\dagger > S$. Observe that $\Pr[D_i \geq \bar{c} | \Gamma_{S^\dagger}] \leq \Pr[D_i \geq \bar{c} | \Gamma_S]$, and $\Pr[D_i \geq \bar{c} | \tilde{\Gamma}_{S^\dagger}] \geq \Pr[D_i \geq \bar{c} | \tilde{\Gamma}_S]$ from which it follows that $\iota(S^\dagger) \leq \iota(S)$. Therefore, the Pragmatic Best Response exists and is unique. It is straightforward to show that the participation rate of pragmatists is increasing in c_G^* .

We turn to proving existence and uniqueness of the Consistent Ethical Response. Let $Q(c)$ denote the fraction of citizens in group G who participate when the ethical cutoff is c ; it follows that $Q(c)$ is increasing in c . By reasoning analogous to Theorem 1, it follows that in a political equilibrium,

$$c_1 Q(c_1) = c_2 Q(c_2),$$

which as before implies that the political equilibrium exists, and the associated ethical cutoff solves $cQ(c) = 2^{1-2\alpha}w/B(\alpha, \alpha)$.