Policing Politicians
Online Appendix

Macartan Humphreys
Columbia University

Jeremy M. Weinstein
Stanford University

September 22, 2012

Abstract

In section 1 we provide a full characterization of the equilibria for the game described in the main paper along with empirical hypotheses derived from the model. In section 2 we provide tables supplementary to results provided in the main text.

1 Model Results and Additional Hypotheses

1.1 Equilibrium

Let $\sigma_{\tilde{s}}$ denote the probability with which the incumbent is returned given signal $\tilde{s}$, and we let $\beta_{\eta\theta}$ denote the probability with which incumbent θ plays $s = 1$ upon observing η. In the second period incumbents play their preferred strategies, as such we focus on strategy choices in the first period only.

We now prove the following proposition which identifies the set of equilibria that can be sustained in each environment.

Proposition 1 The complete set of equilibria are as follows:

[Environment A] If $\tau < -\theta_L$ and $\tau < \theta_H$ then:

- A: There is a unique equilibrium with $\beta_{0L} = \beta_{1H} = 1$, $\beta_{0H} = \beta_{1L} = 0, \sigma_1 = 1, \sigma_0 = 0$.

[Environment B] If $\tau > -\theta_L$ and $\tau < \theta_H$ then:

- B: There are no pure strategy equilibria. In the unique family of mixed strategy equilibria: $\beta_{0H} = 0$, $\beta_{1H} = \beta_{0L} = 1$ and $\beta_{1L} = 2 - \frac{1}{\varphi}$. Voter strategies σ_1, σ_0 are responsive and satisfy $\sigma_1 - \sigma_0 = -\frac{\theta_L}{\tau} \in (0, 1)$.

[Environment C] If $\tau < -\theta_L$ and $\tau > \theta_H$ then:

- $C(i)$ There is a positively responsive pure strategy equilibrium: $\beta_{0L} = 1, \beta_{1L} = 0$, $\beta_{0H} = 1$, $\beta_{1H} = 1, \sigma_1 = 1$, $\sigma_0 = 0$

- $C(ii)$ There is a negatively responsive pure strategy equilibrium: $\beta_{0L} = 1, \beta_{1L} = 0, \beta_{0H} = 0, \beta_{1H} = 0,\ \sigma_1 = 0,\ \sigma_0 = 1$

- $C(iii)$ There is a negatively responsive mixed strategy equilibrium: $\beta_{0L} = 1, \beta_{1L} = 0, \beta_{0H} = 0,\ \beta_{1H} = \frac{1-\varphi}{\varphi}$.Voter strategies $\sigma_1,\ \sigma_0$ satisfy $\sigma_0 - \sigma_1 = \frac{\theta_H}{\tau}$.

[Environment D] If $\tau > -\theta_L$ and $\tau > \theta_H$ then:

- $D(i)$ There is a class of positively responsive pooling equilibria with $\beta_{0L} = \beta_{1H} = \beta_{1L} = \beta_{0H} = 1$. Voter strategies $\sigma_1,\ \sigma_0$ satisfy $\sigma_1 - \sigma_0 > \max(-\frac{\theta_L}{\tau}, \frac{\theta_H}{\tau})$. This class of equilibria includes the pure strategy equilibrium with $\sigma_1 = 1$ and $\sigma_0 = 0$.

- $D(ii)$ There is a class of negatively responsive pooling equilibria with $\beta_{0L} = \beta_{1H} = \beta_{1L} = \beta_{0H} = 0$. Voter strategies $\sigma_1,\ \sigma_0$ satisfy $\sigma_0 - \sigma_1 \geq \max(\frac{-\theta_L}{\tau}, \frac{\theta_H}{\tau}) \in (0,1)$. This class of equilibria includes the pure strategy equilibrium with $\sigma_1 = 0$ and $\sigma_0 = 1$.

- $D(iii)$ If $\theta_H \geq -\theta_L$ there is a class of positively responsive mixed strategy equilibria with $\beta_{0L} = \beta_{1H} = 1,\ \beta_{0H} = 0,\ \beta_{1L} = 2 - \frac{1}{\varphi}$. Voter strategies $\sigma_1,\ \sigma_0$ satisfy $\sigma_1 - \sigma_0 = \frac{-\theta_L}{\tau} \in (0,1)$.

- $D(iv)$ If $\theta_H \geq -\theta_L$ then there is a class of negatively responsive mixed strategy equilibria with: $\beta_{0L} = 1,\ \beta_{1H} = \frac{1}{\varphi} - 1, \beta_{1L} = \beta_{0H} = 0$ and with σ_1 and σ_1 such that $[\sigma_0 - \sigma_1] = \frac{\theta_H}{\tau} \in (0,1)$.

the **Proof** To establish the proposition we first derive a set of relations that hold across environments.

the

Incumbent decision rules If $\eta = 1$, the incumbent will (weakly) prefer $s = 1$ if and only if:

$$\theta + (1-\varepsilon)\sigma_1 + \varepsilon\sigma_0 \geq \varepsilon\sigma_1 + (1-\varepsilon)\sigma_0 \tag{1}$$

Therefore, $\beta_{1\theta} > 0$ only if:

$$\theta \geq (1-2\varepsilon)[\sigma_0 - \sigma_1] = \tau[\sigma_0 - \sigma_1] \tag{2}$$

Similarly, $\beta_{1\theta} < 1$ only if:

$$\theta \leq \tau[\sigma_0 - \sigma_1] \tag{3}$$

If $\eta = 0$, the incumbent will (weakly) prefer $s = 0$ if and only if:

$$\theta + (1-\varepsilon)\sigma_0 + \varepsilon\sigma_1 \geq \varepsilon\sigma_0 + (1-\varepsilon)\sigma_1 \tag{4}$$

Therefore, $\beta_{0\theta} < 1$ only if:

$$\theta \geq -(1-2\varepsilon)[\sigma_0 - \sigma_1] = -\tau[\sigma_0 - \sigma_1] \tag{5}$$

Similarly, $\beta_{0\theta} > 0$ only if:

$$\theta \leq -\tau[\sigma_0 - \sigma_1] \tag{6}$$

These features yield the following relations between voter and incumbent strategies:

$$\begin{aligned}
\theta &> -\tau[\sigma_0 - \sigma_1] \to \beta_{0\theta} = 0 \\
\theta &< -\tau[\sigma_0 - \sigma_1] \to \beta_{0\theta} = 1 \\
\theta &< \tau[\sigma_0 - \sigma_1] \to \beta_{1\theta} = 0 \\
\theta &> \tau[\sigma_0 - \sigma_1] \to \beta_{1\theta} = 1
\end{aligned} \tag{7}$$

The incumbent will be indifferent when $\eta = 1$ if and only if:

$$\theta = \tau[\sigma_0 - \sigma_1]. \tag{8}$$

The incumbent will be indifferent when $\eta = 0$ if and only if:

$$\theta = -\tau[\sigma_0 - \sigma_1] \tag{9}$$

From (8) we have that only one incumbent type can be indifferent if $\eta = 1$, furthermore, a high type can be indifferent only if $\sigma_0 > \sigma_1$ and a low type can be indifferent only if $\sigma_0 < \sigma_1$. From (9) we have that only one type of incumbent can be indifferent if $\eta = 0$, furthermore, a high type can be indifferent only if $\sigma_0 < \sigma_1$ and a low type can be indifferent only if $\sigma_0 > \sigma_1$. Ignoring the possibility that $\theta_H = -\theta_L$ we have that for any pair σ_0, σ_1 only one type can be indifferent and then only in one state. the

Voter Action The voters' decision depends strongly on their posteriors. The voters have a unique best response to return an incumbent if $\tilde{q}(H|\tilde{s} = 1) > q$, and to remove her if $\tilde{q}(H|\tilde{s} = 1) < q$. Mixing is only possible if $\tilde{q}(H|\tilde{s} = 1) = q$. Given strategies $\{\beta_{\eta\theta}\}$, the posterior is given by:

$$\tilde{q}(H|\tilde{s} = 1) = \frac{\Pr(\tilde{s} = 1|\theta = \theta_H, \{\beta_{\eta\theta}\})q}{\Pr(\tilde{s} = 1|\theta = \theta_H, \{\beta_{\eta\theta}\})q + \Pr(\tilde{s} = 1|\theta = \theta_L, \{\beta_{\eta\theta}\})(1-q)}$$

Where:

$$\begin{aligned}
\Pr(\tilde{s} &= 1|\theta = \theta_H, \{\beta_{\eta\theta}\}) = \varphi\left[\beta_{1H}(1-\varepsilon) + (1-\beta_{1H})\varepsilon\right] \\
&\quad + (1-\varphi)\left[\beta_{0H}(1-\varepsilon) + (1-\beta_{0H})\varepsilon\right] \\
\Pr(\tilde{s} &= 1|\theta = \theta_L, \{\beta_{\eta\theta}\}) = \varphi\left[\beta_{1L}(1-\varepsilon) + (1-\beta_{1L})\varepsilon\right] \\
&\quad + (1-\varphi)\left[\beta_{0L}(1-\varepsilon) + (1-\beta_{0L})\varepsilon\right]
\end{aligned}$$

Manipulation of this condition reveals that:

$$\begin{aligned}
\tilde{q}(H|\tilde{s} &= 1) \geq q \leftrightarrow \tilde{q}(H|\tilde{s}=0) \leq q \leftrightarrow \varphi\left(\beta_{1H}-\beta_{1L}\right) \geq (1-\varphi)\left(\beta_{0L}-\beta_{0H}\right) \\
\tilde{q}(H|\tilde{s} &= 1) \leq q \leftrightarrow \tilde{q}(H|\tilde{s}=0) \geq q \leftrightarrow \varphi\left(\beta_{1H}-\beta_{1L}\right) \leq (1-\varphi)\left(\beta_{0L}-\beta_{0H}\right)
\end{aligned} \tag{10}$$

Given these general features we establish the proposition by considering an exhaustive set of cases.

We begin by ruling out equilibria with $\sigma_1 = \sigma_0$, we then identify all "positively responsive" equilibria and finally all "negatively responsive equilibria."

Claim There are no non-responsive equilibria.

Assume contrary to the claim that $\sigma_1 = \sigma_0$ in equilibrium.

Recall that $\beta_{1\theta} = 0$ if $\theta < \tau[\sigma_0 - \sigma_1] = 0$ and $\beta_{1\theta} = 1$ if $\theta > \tau[\sigma_0 - \sigma_1] = 0$.We then have: $\beta_{1L} = 0$, $\beta_{1H} = 1$. Since $\beta_{0\theta} = 0$ if $\theta > -\tau[\sigma_0 - \sigma_1] = 0$ and $\beta_{0\theta} = 1$ if $\theta < -\tau[\sigma_0 - \sigma_1] = 0$, and therefore $\beta_{0L} = 1$, $\beta_{0H} = 0$.

Given these strategies we have:

$$\tilde{q}(H|\tilde{s}=1) > q \leftrightarrow \varphi\left(\beta_{1H}-\beta_{1L}\right) > (1-\varphi)\left(\beta_{0L}-\beta_{0H}\right) \leftrightarrow \varphi > \frac{1}{2}$$

However $\varphi > \frac{1}{2}$ by assumption and so $\tilde{q}(H|\tilde{s}=1) > q$ which implies $\sigma_1 = 1$ in equilibrium. Similarly $\tilde{q}(H|\tilde{s}=0) < q$ which requires $\sigma_1 = 0$.

Claim Environment A: There is a unique equilibrium

In environment A, from (7) we have $\theta_L < -\tau$ implies $\beta_{0L} = 1$ and $\beta_{1L} = 0$ and $\theta_H > \tau$ implies $\beta_{0H} = 0$ and $\beta_{1H} = 1$.

The unique equilibrium involves pure strategies in which H plays good policies and L chooses bad policies. Voters infer that an incumbent is of a high type if and only if they observe $\tilde{s} = 1$.

Claim Environment B: There are no Pure Strategy Equilibria. There is a single class of Mixed Strategy Equilibria.

Consider first a positively responsive pure strategy with $[\sigma_0 - \sigma_1] = -1$. Then, from (7): $\beta_{1\theta} = 0$ if $\theta < -\tau$, $\beta_{1\theta} = 1$ if $\theta > -\tau$, $\beta_{0\theta} = 0$ if $\theta > \tau$ and $\beta_{0\theta} = 1$ if $\theta < \tau$.Any such equilibrium must involve $\beta_{0H} = 0$ and $\beta_{1H} = \beta_{1L} = \beta_{0L} = 1$. In this case $\tilde{q}(H|\tilde{s}=1) < q \leftrightarrow \varphi\left(\beta_{1H}-\beta_{1L}\right) < (1-\varphi)\left(\beta_{0L}-\beta_{0H}\right) \leftrightarrow 0 < (1-\varphi)$. Hence if the voter observes a $\tilde{s} = 1$ she will infer that the incumbent is more likely to be of type L and remove her, contrary to the assumption.

Consider next a negatively responsive pure strategy with $[\sigma_0 - \sigma_1] = 1$. Then: $\beta_{1\theta} = 0$ if $\theta < \tau$, $\beta_{1\theta} = 1$ if $\theta > \tau$, $\beta_{0\theta} = 0$ if $\theta > -\tau$ and $\beta_{0\theta} = 1$ if $\theta < -\tau$ and so in equilibrium we require: $\beta_{1H} = 1, \beta_{0H} = \beta_{1L} = \beta_{0L} = 0$.

In this case $\tilde{q}(H|\tilde{s}=1) > q \leftrightarrow \varphi\left(\beta_{1H}-\beta_{1L}\right) > (1-\varphi)\left(\beta_{0L}-\beta_{0H}\right) \leftrightarrow \varphi > 0$ and so if the voter observes a $\tilde{s} = 1$ she will infer that the incumbent is more likely to be of type H and retain him, contrary to the assumption.

Hence the only equilibria in Environment B are mixed strategy equilibria.

In a mixed strategy equilibrium the requirement for the incumbent to mix is: $\theta = \tau[\sigma_0 - \sigma_1]$ if $\eta = 1$ and $\theta = -\tau[\sigma_0 - \sigma_1]$ if $\eta = 0$.

Condition $\theta_H > \tau$ together with the fact that $\sigma_0 - \sigma_1 \leq 1$ implies that H will never mix and in particular $\beta_{0H} = 0$ and $\beta_{1H} = 1$. Hence any mixing must be by L only. When $\eta = 1$, we need $[\sigma_1 - \sigma_0] = -\frac{\theta_L}{\tau}$ where $0 \leq -\frac{\theta_L}{\tau} \leq 1$. When $\eta = 0$, we need $[\sigma_0 - \sigma_1] = -\frac{\theta_L}{\tau}$, where again $0 \leq -\frac{\theta_L}{\tau} \leq 1$. Thus a σ_0, σ_1 combination can be chosen in which L will mix under one but only one state of the world.

We examine each case. Assume first that $\beta_{0L} = 1$. Then:

$$\tilde{q}(H|\tilde{s} = 1) = q \leftrightarrow \varphi\left(\beta_{1H} - \beta_{1L}\right) = (1 - \varphi)\left(\beta_{0L} - \beta_{0H}\right) \leftrightarrow \beta_{1L} = 2 - \frac{1}{\varphi}$$

If however $\tilde{s} = 0$ then:

$$\tilde{q}(H|\tilde{s} = 0) = q \leftrightarrow \beta_{1L} = 2 - \frac{1}{\varphi}$$

Hence with $\beta_{1L} = 1$, mixing can be sustained either when $\tilde{s} = 0$ or $\tilde{s} = 1$ or both.

Assume next that $\beta_{1L} = 0$. Then:

$$\tilde{q}(H|\tilde{s} = 1) = q \leftrightarrow \varphi\left(\beta_{1H} - \beta_{1L}\right) = (1 - \varphi)\left(\beta_{0L} - \beta_{0H}\right) \leftrightarrow \beta_{0L} = \frac{\varphi}{1 - \varphi} > 1$$

Similarly:

$$\tilde{q}(H|\tilde{s} = 0) = q \leftrightarrow \beta_{0L} = \frac{\varphi}{1 - \varphi} > 1$$

And no mixing cannot be sustained. Thus the unique family of mixed strategy equilibria involve strategies: $\beta_{0H} = 0$, $\beta_{1H} = \beta_{0L} = 1$ and $\beta_{1L} = 2 - \frac{1}{\varphi}$. The voters have a set of feasible strategies over σ_1, σ_0 such that $[\sigma_1 - \sigma_0] = -\frac{\theta_L}{\tau}$ and hence $\sigma_1 > \sigma_0$.

Claim Environment C: Positively Responsive Equilibria Imply Pure Strategies.

Assume that in equilibrium: $\sigma_1 > \sigma_0$. Then from (7) we have: $\beta_{0L} = 1$ and $\beta_{1H} = 1$.

From $\theta_L < -\tau$ we have $\theta_L < -\tau[\sigma_1 - \sigma_0] < \tau[\sigma_1 - \sigma_0]$ and so $\beta_{1L} = 0$. Adding these elements together we have:

$$\tilde{q}(H|\tilde{s} = 1) > q \leftrightarrow \varphi > (1 - \varphi)\left(1 - \beta_{0H}\right)$$

Thus for all values of β_{0H} we have $\tilde{q}(H|\tilde{s} = 1) > q$ and hence there is no mixed strategy equilibrium, and in particular, $\sigma_1 = 1$ and $\sigma_0 = 0$. Using this fact we have that in the unique responsive equilibrium in environment C, $\theta_H < \tau$ implies $\beta_{0H} = 1$.

Claim Environment D: Positively Responsive Equilibria are all of Type $D(i)$ or $D(iii)$

If $\sigma_1 > \sigma_0$ then, from (7) we have : $\beta_{0L} = 1$ and $\beta_{1H} = 1$. In this case:

$$\tilde{q}(H|\tilde{s} = 1) \geq q \leftrightarrow \varphi\left(1 - \beta_{1L}\right) \geq (1 - \varphi)\left(1 - \beta_{0H}\right)$$

For a responsive pure strategy equilibrium we have $\sigma_1 = 1$ and $\sigma_0 = 0$ and so, $\beta_{0H} = 1$ and $\beta_{1L} = 1$.

For mixing to be possible in a positively responsive equilibrium we require $\varphi\left(1 - \beta_{1L}\right) = (1 - \varphi)\left(1 - \beta_{0H}\right)$ and either (i) $\beta_{1L} = \beta_{0H} = 1$ or (ii) $\left(1 - \beta_{1L}\right) < \left(1 - \beta_{0H}\right)$ and so $\beta_{0H} < \beta_{1L}$.

For (i) we need (for $\beta_{1L}=1$) that $\theta_L > \tau[\sigma_0-\sigma_1]$ and (for $\beta_{0H}=1$) that $\theta_H < -\tau[\sigma_0-\sigma_1]$. For this we need: $\sigma_1-\sigma_0 > \max(-\frac{\theta_L}{\tau},\frac{\theta_H}{\tau})$. This class of equilibria ($D(i)$) includes the pure strategy equilibrium with $\sigma_1=1$ and $\sigma_0=0$.

For case (ii) $\beta_{0H}<\beta_{1L}$ implies that $\beta_{0H}<1$ and $\beta_{1L}>0$. We have established that it is not possible for both types to mix in any equilibrium, furthermore we can rule out the possibility that H mixes since in that case $\beta_{1L}=1$, but then the condition $(1-\beta_{1L})<(1-\beta_{0H})$ cannot be satisfied. The only mixing then involves L mixing, and so $\beta_{0H}=0$ and $\beta_{1L}=2-\frac{1}{\varphi}$. To support this equilibrium we require that $\theta_L=\tau[\sigma_0-\sigma_1]$ and so $[\sigma_1-\sigma_0]=\frac{-\theta_L}{\tau}$. In addition to support $\beta_{0H}=0$ we need, from 5, that $\theta_H \geq -\tau[\sigma_0-\sigma_1]=-\theta_L$. This is case $D(iii)$.

Claim The only negatively responsive equilibrium are those given by $C(ii)$, $C(iii)$, $D(ii)$ and $D(iv)$.

Assume that in equilibrium: $\sigma_1<\sigma_0$.

If $\eta=1$, the incumbent will prefer to play $s=1$ if and only if: $\theta \geq \tau[\sigma_0-\sigma_1]$. With $\sigma_1<\sigma_0$, the low type will always play $s=0$ if $\eta=1$, that is: $\beta_{1L}=0$.

If $\eta=0$, the incumbent will prefer to play $s=0$ if and only if: $\theta \geq -\tau[\sigma_0-\sigma_1]>0$. With $\sigma_1<\sigma_0$, the high type will always play $s=0$ if $\eta=0$. That is: $\beta_{0H}=0$.

To sustain $\sigma_1<\sigma_0\leq 1$ we require $\tilde{q}(H|\tilde{s}=1)\leq q$, or equivalently:

$$\begin{aligned}\tilde{q}(H|\tilde{s} &= 1)\leq q \leftrightarrow \varphi\left(\beta_{1H}-\beta_{1L}\right)\leq(1-\varphi)\left(\beta_{0L}-\beta_{0H}\right)\\ &\leftrightarrow \varphi\beta_{1H}\leq(1-\varphi)\beta_{0L}\end{aligned}$$

Thus (since $\varphi>.5$,) we require that either (i) $\beta_{1H}=\beta_{0L}=0$ or (ii) $\beta_{1H}<\beta_{0L}$ and in particular that $\beta_{0L}>0$ and $\beta_{1H}<1$.

In case (i) $\beta_{1H}=\beta_{0L}=0$ requires that (a) $\theta_H\leq\tau[\sigma_0-\sigma_1]$ and (b) $\theta_L\geq-\tau[\sigma_0-\sigma_1]$. This can only be sustained in environment D. To see this note that condition (a) can never be satisfied if $\theta_H>\tau$ and this allows us to rule out negatively responsive equilibria in environments A and B. Condition (b) can never be satisfied if $\theta_L<-\tau$ or $-\theta_L>\tau$ and this allows us to rule out environment C. In environment D however pooling of this form is possible if $\sigma_0-\sigma_1\geq\max(\frac{-\theta_L}{\tau},\frac{\theta_H}{\tau})$. This corresponds to case $D(ii)$.

The conditions in case (ii) themselves imply that: $\theta_H\leq\tau[\sigma_0-\sigma_1]$ and $\theta_L\leq-\tau[\sigma_0-\sigma_1]$ or $-\theta_L\geq\tau[\sigma_0-\sigma_1]$. The condition $\theta_H\leq\tau[\sigma_0-\sigma_1]$ can never be satisfied if $\theta_H>\tau$ and this allows us to rule out negatively responsive equilibria in environments A and B. Together these imply that $\theta_H\leq-\theta_L$ which holds in case C.

A negatively responsive pure strategy equilibrium in case (ii) thus requires $\beta_{0L}=1$ and $\beta_{1H}=0$. No such equilibrium holds in environment D since for $\beta_{0L}=1$ we require $\theta_L\leq -\tau[\sigma_0-\sigma_1]=-\tau$ which holds only in environments A and C. We have already rules out such an equilibrium in environment A; such an equilibrium does obtain in environment C however and corresponds with equilibrium $C(ii)$.

A negatively responsive mixed strategy equilibrium in environment C can only be sustained if $\varphi\beta_{1H}=(1-\varphi)\beta_{0L}$. Since mixing can only take place with respect to one strategy we need $\beta_{0L}=1$ and $\beta_{1H}=\frac{1-\varphi}{\varphi}\in(0,1)$ (note $\beta_{1H}=1$ implies $\beta_{0L}=\frac{\varphi}{1-\varphi}>1$) and $\sigma_0-\sigma_1=\frac{\theta_H}{\tau}$. This corresponds to equilibrium $C(iii)$. Note that to sustain $\beta_{0L}=1$ we need $\theta_L<-\tau[\sigma_0-\sigma_1]=-\theta_H$ which is true in environment C.

A negatively responsive equilibrium in environment D can only be sustained if $\tau > -\theta_L \geq \tau[\sigma_0 - \sigma_1]$, and hence if $[\sigma_0 - \sigma_1] < 1$. Equivalently, to sustain a negatively responsive equilibrium in environment D, some voter type must mix. However mixing requires that in equilibrium $\tilde{q}(H|\tilde{s} = 1) = q$, and so $\frac{\varphi}{1-\varphi}\beta_{1H} = \beta_{0L}$. Since $\beta_{1H} < \beta_{0L}$ this condition cannot be met be $\beta_{1H} = \beta_{0L} = 0$, instead mixing by one or other incumbent type is required. In addition the condition cannot be met if $\beta_{1H} = 0$ or $\beta_{0L} = 0$. Therefore we have $\beta_{1H} > 0$ and $\beta_{0L} > 0$. Generically we have established that only one type will mix for a given voter strategy. Since $\frac{\varphi}{1-\varphi} > 1$, the only feasible mixed strategy equilibrium requires $\beta_{1H} = \frac{1-\varphi}{\varphi}, \beta_{0L} = 1$. H will be *willing* to mix iff $\theta_H = \tau[\sigma_0 - \sigma_1]$, that is: $[\sigma_0 - \sigma_1] = \frac{\theta_H}{\tau}$. And, from 5, L will be willing to play $\beta_{0L} = 1$ only if $\theta_L \geq -\tau[\sigma_0 - \sigma_1] = -\theta_H$. This corresponds to case $D(iv)$.

1.2 Implications

1.2.1 Welfare Implications

Consider now the question of voter welfare. Total expected voter utility in environment A is given as follows:

$$\begin{aligned} W(A) &= q[1 + \varphi[(1-\varepsilon) + \varepsilon q] + (1-\varphi)[\varepsilon + (1-\varepsilon)q]] \\ &\quad +(1-q)[\varphi(1-\varepsilon) + (1-\varphi)\varepsilon]q \\ &= q[1+q] + 2q(1-q)[\varphi + (1-2\varphi)\varepsilon] \end{aligned}$$

We can see from this equation that welfare is increasing in transparency within equilibria of type A; in addition, the gains from transparency are greatest when prior uncertainty about the incumbent types is high ($q = .5$) and uncertainty about the correct type of policy is low ($\varphi = 1$). In environment B we have:

$$W(B|q, \varphi, \varepsilon) = (2\varphi - 1) + q(3 - 2\varphi) - (1-\varphi)^2(1-q)q\varepsilon$$

Within environment B, the gains from transparency are greatest when prior uncertainty about the incumbent types is high ($q = .5$) and uncertainty about the correct type of policy is high ($\varphi = .5$); but even in these cases the marginal effect is much weaker than in environment A.

Welfare in environments C and D are more straightforward:

$$\begin{aligned} W(C|q, \varphi, \varepsilon) &= q\left[1 + (2-q)\varphi\right] - 2\varphi q(1-q)\varepsilon \\ W(D|q, \varphi, \varepsilon) &= \varphi + q \end{aligned}$$

In all four environments it is easy to check that $\frac{\partial W}{\partial \varepsilon} \leq 0$, with the inequality strict for all but case D. This implies that, *locally,* transparency produces gains in welfare; these *local* gains are due entirely to a better ability to select MPs. However the effects of accountability mechanism are more complicated: a rise in transparency can be associated with a fall in voter welfare if the equilibrium shifts from one environment to another. Indeed this is the key result of the analysis: *globally, a rise in transparency can have positive, negative or non-monotonic effects depending on the underlying parameter values.*

Figure 1 shows how welfare depends on transparency for a range of parameter values. The three left graphs consider cases in which environments A, B and D obtain. Specifically we impose $\theta_H = \frac{2}{3}$, $\theta_L = -\frac{1}{3}$. The right three graphs show equilibria in environments A, C and D for a case with $\theta_H = \frac{1}{3}$ and $\theta_L = -\frac{2}{3}$. Each graph considers a different value for φ, as marked on the titles, and within each graph the four lines correspond (in order from bottom to top) to $q = 0$, $q =, 5$, $q = .75$ and $q = 1$.

The lower figures correspond to cases in which $\varphi = 1$ (in which there is no difficulty in associating good actions with good outcomes). In these cases the more transparency the better. Within environment A, more transparency leads to better selection of second stage politicians, and thus a rise in welfare. The major gains arise however from shifts from environments A to B and from C to D. These step shifts are pure accountability shifts; they correspond exactly to the gains from inducing bad types to take action $s = 1$.

The central panels ($\varphi = .75$) in which there is a positive but imperfect relation between actions and outcomes, tell a more complex story. In some cases, a rise in transparency leads to a rise in welfare throughout its range. This is true for example if almost all types are Low, $q \approx 0$. However in other cases, notably when $q \approx 1$ transparency has the opposite, adverse effect. In these cases, the (many) good types who would select policies they know to be good under equilibria A or B choose instead to conform, knowing that whenever $\eta = 0$, their good actions run a risk of being misinterpreted by voters. As a consequence, they conform to expectations instead of seeking to achieve public benefits. In intermediate cases, non-monotonicities can arise, with a rise in transparency leading to either an intermediate rise or decline in welfare. *Which type of non-monotonicity arises depends on the relative gains from incentivizing bad types to act well when* $\eta = 1$ *and the losses associated with good types acting badly when* $\eta = 0$.

Finally, we note that even when $\varphi \approx 0.5$ and there is no (ex ante) relationship between s and benefits to voters, the first column in Figure 1 tells us that information about s nevertheless can help keep politicians accountable. In the extreme case of *only bad politicians*, a rise in transparency allows voters to ensure that politicians choose the right action half the time (although voters never know which half); the same adverse effects seen in the $\varphi = .75$ cases do however obtain here also.

From these observations we derive the following hypothesis:

$H_{Welfare}$ (Welfare Gains) A rise in transparency is associated with gains in voter welfare in cases in which MPs are not believed to have voter interests at heart and in which voters are more confident of the mapping between actions and outcomes, but is associated with a fall in welfare when MPs are believed to have voter interests at heart or in which voters are less confident of the mapping between actions and outcomes.

In the case in which only the selection mechanism is in operation, there are unambiguous gains in the second period and no effects on welfare in the first period. In the case in which only the accountability mechanism is in operation, there are ambiguous effects in the first period and no effects on welfare in the second period.

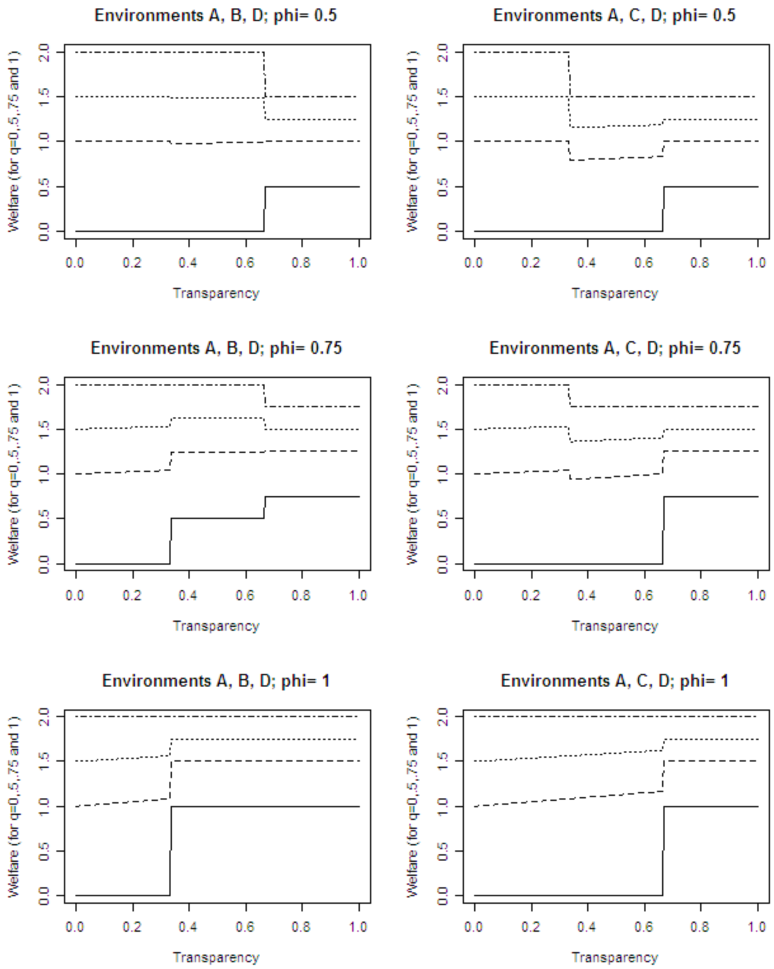

Figure 1: Citizen welfare as a function of τ for a series of parameter values. In each graph higher lines correspond to higher values of q.

1.2.2 Reelection Probabilities

As shown in Figure 2, there is a non-monotonic relationship between transparency and turnover. In all cases if transparency is already sufficiently high as to ensure good performance through the accountability mechanism, a rise in transparency *reduces* turnover rates by ensuring that voters are less likely to make false judgments. However transparency can also increase turnover through a number of channels. In environment A, for example, if politicians are implementing their preferred strategies, unrestrained by voters, a rise in transparency can still facilitate selection by reducing the likelihood of removing High types and increasing the likelihood of removing Low types.

General hypotheses are hard to draw and again depend on beliefs about the incumbents types and confidence in policy mappings. We extract the following, however, for study:

$H_{Incumbency}$ (Incumbency Advantage) The incumbency advantage is increasing in transparency when

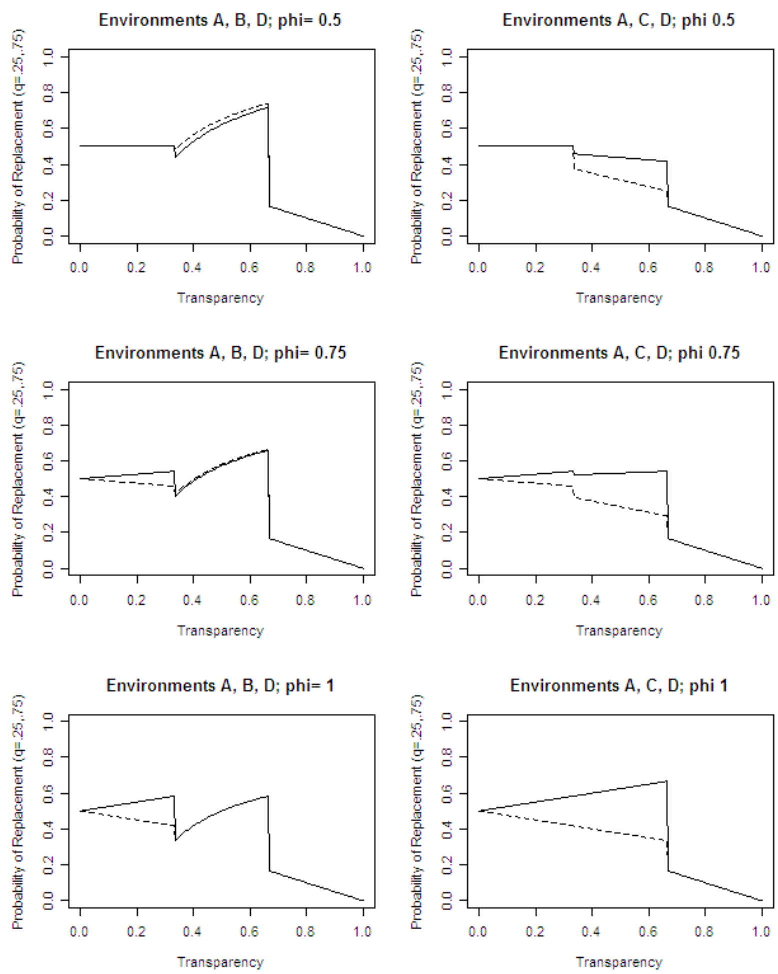

Figure 2: Probability with which the incumbent is replaced as a function of τ for a series of parameter values given $q = .25$ (solid line) and $q = .75$ (dotted line).

there is greater uncertainty regarding the mapping from action to outcomes or when the prior pool of candidates is believed to have voter interests at heart. Turnover rates are increasing at intermediate levels of transparency, especially when there is prior distrust in politicians and when the mapping from inputs to outputs is believed to be known.

1.2.3 Interaction between selection and accountability mechanisms

Let $\overline{q}^X$ denote the probability with which a a replaced incumbent is in fact of high quality in environment X. Then the probability that the replacement is higher quality than the replaced incumbent is $q(1 - \overline{q}^X)$; selection then results in higher quality politicians when q^X is low.

It is easy to see however that $\overline{q}^D > \overline{q}^A$. Similarly $\overline{q}^D > \overline{q}^C$. But $\overline{q}^A > \overline{q}^C$.These results suggest that in some ranges the accountability mechanism inhibits the selection mechanism from operating, but in other ranges it enhances the selection mechanism. Increased transparency that results in greater conformism by high types (but not low types) improves selection (the shift from environment A to environment B). But improvements in transparency that result in greater conformism by low types weakens selection (the shift from environment C to environment D). The overall effect over the full range (A to D) is a worsening in selection. This contrasts sharply with what would arise in a situation in which incumbents are not responsive to electoral incentives. In these environments an increase in transparency unambiguously improves expected politician quality in all ranges.

$H_{Selection}$ (Quality of replacements) Greater transparency results in weakened expectations that replacement politicians are of higher quality than incumbents.

Note that if only the selection mechanism is in operation, then we expect the opposite relationship of that given in the hypothesis to hold.

More formally:

$$\begin{aligned}
\overline{q}^A &= \frac{\kappa q}{\kappa q + (1-\kappa)(1-q)} \\
\overline{q}^B &= \frac{\left(\kappa + (1-\kappa)\frac{\tau+\theta_L}{\tau}\right) q}{\left(\kappa + (1-\kappa)\frac{\tau+\theta_L}{\tau}\right) q + (((2-\frac{1}{\varphi})((1-\epsilon)\frac{\tau+\theta_L}{\tau} + \epsilon) + (\frac{1}{\varphi}-1)(\kappa\frac{\tau+\theta_L}{\tau} + (1-\kappa)))(1-q)} \\
\overline{q}^C &= \frac{\epsilon q}{\epsilon q + (1-\kappa)(1-q)} \\
\overline{q}^D &= \frac{\epsilon q}{\epsilon q + \epsilon(1-q)} = q
\end{aligned} \tag{11}$$

where $\kappa \equiv \varphi\epsilon + (1-\varphi)(1-\epsilon)$ is the probability that voters will observe a signal of bad policies given a politician is implementing bad policies. It is also the probability that voters will observe a signal of good policies when bad policies are being played. Note that $1-\kappa = \varphi(1-\epsilon) + (1-\varphi)\epsilon = (1-\varphi)\epsilon + \varphi(1-\epsilon)$ which is the probability that someone playing bad policies will be seen as playing bad policies and that someone playing good policies will be seen as playing good policies.

It follows that $\overline{q}^D > \overline{q}^C$ and $\overline{q}^D > \overline{q}^A$ but $\overline{q}^A > \overline{q}^C$.

For: $\overline{q}^D > \overline{q}^C$ note:

$$\overline{q}^D > \overline{q}^C \leftrightarrow \epsilon < 1-\kappa \leftrightarrow \epsilon < \varphi\epsilon + (1-\varphi)(1-\epsilon) \leftrightarrow \epsilon < .5$$

For $\overline{q}^C < \overline{q}^A$:

$$\overline{q}^C < \overline{q}^A \leftrightarrow \frac{\epsilon q}{\epsilon q + (1-\kappa)(1-q)} < \frac{\kappa q}{\kappa q + (1-\kappa)(1-q)}$$
$$\leftrightarrow$$
$$(\kappa q + (1-\kappa)(1-q))\epsilon q < (\epsilon q + (1-\kappa)(1-q))\kappa q \leftrightarrow \epsilon < \kappa \leftrightarrow \epsilon < .5$$

For $\overline{q}^A < \overline{q}^D$:

$$\overline{q}^A < \overline{q}^D \leftrightarrow \frac{\kappa q}{\kappa q + (1-\kappa)(1-q)} < q \leftrightarrow \kappa < .5$$

which always holds under our assumptions of $\varphi > .5$ and $\epsilon < .5$.

1.2.4 Candidate Pool

Finally we can consider the incentives for individuals to stand as MPs for any given level of transparency. We suppose again that $|\theta_i| < 1$ and hence that the maximum utility obtainable from office is less than 2. Finally we assume that there are an equal number of good and bad potential candidate types and that each individual has an outside option distributed $u \sim U[0,2]$. We expect that candidates will stand for office only if their expected gains, $y > u$.

Our interest is in determining whether the composition of the candidate pool is likely to improve or worsen with transparency.

The expected benefit to a candidate of type H in equilibrium A is:

$$u_{HA} = \theta_H + \varphi(1-\varepsilon) + (1-\varphi)\varepsilon = \theta_H + \varphi + (1-2\varphi)\varepsilon$$

To place the utilities of the High and Low types on a comparable scale (relative to u) we add an extra term $-\theta_L$ to the Low types utility. The expected benefit to a candidate of type L in an equilibrium in environment A is then:

$$u_{LA} = -\theta_L + \varphi\varepsilon + (1-\varphi)(1-\varepsilon) = -\theta_L - (1-2\varphi)\varepsilon + (1-\varphi)$$

The share of candidates that are high types from the pool of candidates willing to stand for office at the beginning of the first period is then simply:

$$q_A = \frac{u_{HA}\frac{1}{2}}{u_{HA}\frac{1}{2} + u_{LA}\frac{1}{2}} = \frac{\theta_H + \varphi + (1-2\varphi)\varepsilon}{\theta_H - \theta_L + 1}$$

which is decreasing in ε. Hence more transparency produces a better pool.

In a similar way we have:

$$\begin{aligned}
q_B &= \frac{\theta_H + (\varphi + (1-2\varphi)\varepsilon)\frac{-\theta_L}{\tau}}{\theta_H - \theta_L + \frac{-\theta_L}{\tau}} \\
q_C &= \frac{\varphi\theta_H + (1-\varepsilon)}{\varphi\theta_H - \theta_L + 1 + (1-\varphi)(1-2\varepsilon)} \\
q_D &= \frac{\varphi\theta_H + (1-\varepsilon)}{\varphi\theta_H - (1-\varphi)\theta_L + 2(1-\varepsilon)}
\end{aligned}$$

From these values we can establish that q_A, q_B and q_C are decreasing in ε. However, q_D can be increasing or decreasing in ε depending on whether office is a more attractive prospect for high or low types. It is increasing in ε if and only if: $\frac{\theta_H}{-\theta_L} > \frac{1-\varphi}{\varphi}$ and decreasing if and only if $\frac{\theta_H}{-\theta_L} < \frac{1-\varphi}{\varphi}$. Hence q_D will be increasing in ε (that is, falling in transparency) whenever $\theta_H > -\theta_L$ and whenever the mapping from outcomes is well known (φ close to 1).

As before, a change in ε can also be associated with a change in the type of equilibrium, with more dramatic consequences for behavior. Note that if $\frac{-\theta_L}{\tau} = 1$, then: $q_B = q_A$; this establishes that the share of H types is increasing over the range between equilibria type A and equilibria type B. Similarly when $\frac{\theta_H}{\tau} = 1$, $q_C = q_A$ which establishes that the gain from transparency holds across these parameter ranges also. Hence the pool of candidates is improving in transparency in low and intermediate ranges.[1]

However, in ranges in which players are already pooling on conformist action, or in which a rise in transparency induces them to pool, rising transparency has adverse effects on the pool of applicants. The between-environment fall in the quality of the candidate pool for a shift from state B to D arises from two effects: from the fact that High types now conform in order to ensure reelection, and from the fact that Low types, though willing to conform in equilibrium B, are more likely to be rewarded for conforming in equilibrium D. The intuition for the worsening pool of candidates *within* equilibrium D is the following. Each type's benefit comes from two elements — the Period 1 benefit, which is greater for the High type than for the Low type, and the period 2 benefit, which is equal across both types. As transparency rises, the expected gains to both types of Period 2 benefits rises and in doing so it reduces the relative aggregate gains of High types compared to Low types.

H_{Pool} (Candidate pool) A rise in transparency will be associated with an improvement in the quality of the pool of candidates (and, relative to the control areas, a larger positive difference between the performance of newly elected MPs after the 2011 elections and that of the candidates that they replaced), at low levels of transparency, with this effect weakening or reversing at high levels of transparency.

[1]In addition, we have that provided $\frac{\theta_H - \theta_L\theta_H + \frac{1}{2}\theta_L}{\frac{1}{2}\theta_H - 2\theta_L\theta_H - \frac{1}{2}\theta_L} < \varphi$, the pool contains relatively more high types in the full transparency state ($\tau = 1$) than in the lowest transparency state ($\tau = 0$); this condition always holds with $\theta_H < -\theta_L$ (that is when the relevant environments are A, C, D) and can never hold if $\frac{\theta_H}{2\theta_H+2} > -\theta_L$. For $\frac{\theta_H}{2\theta_H+2} < -\theta_L < \theta_H$ improvements in the pool across the full range depend on the quality of the signal φ.

2 Extra Tables

	Elect	Elect	Elect	Ran	Ran	Ran	Share	Share	Share
Workshop (2sls)	-0.031	-0.084	0.071	-0.085	-0.054	-0.009	0.032	-0.078	-0.019
	(0.20)	(0.52)	(0.47)	(0.65)	(0.39)	(0.07)	(0.48)	(1.08)	(0.31)
Interaction (2sls)	0.000	0.001	-0.002	0.002	0.001	-0.000	-0.001	0.001	0.000
	(0.07)	(0.39)	(0.78)	(0.77)	(0.38)	(0.02)	(0.62)	(0.90)	(0.26)
Plenary pct	-0.000			0.000			-0.000		
	(0.08)			(0.27)			(0.61)		
Committee pct		-0.000			0.001			-0.001	
		(0.24)			(0.45)			(0.98)	
Constituency pct			0.004			0.002			0.001
			(3.07)***			(1.35)			(2.26)**
Constant	0.462	0.464	0.242	0.758	0.746	0.700	0.478	0.486	0.385
	(5.32)	(4.78)	(2.98)	(10.42)	(9.11)	(10.13)	(12.30)	(10.80)	(10.49)
R^2	.	0.00	0.04	0.00	0.00	0.01	0.01	0.01	0.05
N	293	239	292	293	239	292	227	185	227

Table 1: Local Average Treatment Effects of Dissemination Workshops. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

	Piped Water	Primary Educa- tion	Education Facility	Teacher Presence	Health Facility	Doctor Present	Perceptions Index
Workshop	-0.036	-0.028	-0.048	0.029	0.022	-0.073	0.018
	(0.62)	(0.73)	(1.10)	(1.16)	(0.63)	(2.16)**	(0.19)
Constant	0.602	0.622	0.379	0.821	0.309	0.390	2.295
	(13.82)	(25.47)	(12.48)	(41.43)	(12.26)	(19.55)	(31.33)
R^2	0.00	0.00	0.00	0.00	0.00	0.01	0.00
N	2,054	1,934	1,178	973	2,102	2,118	2,034

Table 2: Adverse Effects at the District Level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the district level.