A Supplementary Material

A.1 The probability of being voted [Eq. (9)]

Suppose that parties have already chosen the platform vector (y_A, y_B) while Nature has already sent signals $\left(q_A^i, q_B^i\right)$. The representative voter i votes for party A if

$$z \geq \widehat{z}(y_A, q_A^i, y_B, q_B^i) \qquad \widehat{z}(y_A, q_A^i, y_B, , q_B^i) \equiv E\left(U_B | y_B, q_B^i\right) - E\left(U_A | y_A, q_A^i\right)$$

Thus, from the perspective of party A, the probability of being voted by the representative voter is

$$E_{q_A^i, y_B, q_B^i}\left[\Pr\left(z \geq \widehat{z}(y_A, q_A^i, y_B, q_B^i)\right)\right] =$$

$$= E_{q_A^i, y_B, q_B^i}\left[\frac{1}{2} - \frac{\widehat{z}(y_A, q_A^i, y_B, q_B^i)}{\overline{z}}\right]$$

$$= \frac{1}{2} + \frac{1}{\overline{z}}\left[E_{q_A^i} E\left(U_A | y_A, q_A^i\right) - E_{y_B, q_B^i} E\left(U_B | y_B, q_B^i\right)\right] \tag{A.1.1}$$

Eq.(A.1.1) coincides with Equation (9) in the main text. To derive Eq.(A.1.1) we have assumed that $\eta_j = \eta + \widetilde{\eta}_j$ with η distributed over an unbounded support and $\widetilde{\eta}_j$ distributed over a bounded support and independent form $\widetilde{\eta}_{j'}$. This implies that the support of $\widehat{z}(y_A, q_A^i, y_B, , q_B^i)$ is bounded too. Further, in line with Matejika and Tabellini (2020), we have also assumed that the probability of $\widehat{z}(y_A, q_A^i, y_B, , q_B^i)$ falling outside the support of z is negligible. This requires $\overline{z}$ to be sufficiently large.

A.2 The properties of $y(\eta)$

In this section, we establish that the function $y(\eta)$ is continuous, monotone and differentiable.

Continuity

Argument: by contradiction. Suppose that $y(\eta)$ is not continuous at some η_0 belonging to the support of η. Since $y(\eta)$ maximizes the objective (4), for a party endowed with competence $\eta_0 - \varepsilon$ $[\eta_0 + \varepsilon]$ it is optimal to set $y(\eta_0 - \varepsilon)$ $[\, y(\eta_0 + \varepsilon)]$, with ε representing a small positive amount:

$$\rho P(y(\eta_0 - \varepsilon), \eta_0 - \varepsilon) + W(y(\eta_0 - \varepsilon), \eta_0 - \varepsilon) \geq \rho P(y(\eta_0 + \varepsilon), \eta_0 - \varepsilon) + W(y(\eta_0 + \varepsilon), \eta_0 - \varepsilon) \tag{A.2.1}$$

$$\rho P(y(\eta_0 + \varepsilon), \eta_0 + \varepsilon) + W(y(\eta_0 + \varepsilon), \eta_0 + \varepsilon) \geq \rho P(y(\eta_0 - \varepsilon), \eta_0 + \varepsilon) + W(y(\eta_0 - \varepsilon), \eta_0 + \varepsilon) \tag{A.2.2}$$

By letting $\varepsilon \longrightarrow 0$, and by using the expressions for P and W, Inequalities (A.2.1) and (A.2.2) can be written respectively as follows:

$$\lim_{\varepsilon\longrightarrow 0}\left\{g(\eta_0-\varepsilon)-\frac{1}{2}\left[-g(\eta_0-\varepsilon)+\eta_0-T\right]^2\right\}\geq\lim_{\varepsilon\longrightarrow 0}\left\{g(\eta_0+\varepsilon)-\frac{1}{2}\left[-g(\eta_0+\varepsilon)+\eta_0-T\right]^2\right\} \tag{A.2.3}$$

$$\lim_{\varepsilon\longrightarrow 0}\left\{g(\eta_0+\varepsilon)-\frac{1}{2}\left[-g(\eta_0+\varepsilon)+\eta_0-T\right]^2\right\}\geq\lim_{\varepsilon\longrightarrow 0}\left\{g(\eta_0-\varepsilon)-\frac{1}{2}\left[-g(\eta_0-\varepsilon)+\eta_0-T\right]^2\right\} \tag{A.2.4}$$

Inequalities (A.2.3) and (A.2.4) are mutually consistent only if

$$\lim_{\varepsilon\longrightarrow 0} y(\eta_0-\varepsilon)=\lim_{\varepsilon\longrightarrow 0} y(\eta_0+\varepsilon) \tag{A.2.5}$$

that is, in contradiction with the premise, if $y(\eta)$ is continuous at η_0.◇

Monotonicity

Argument: monotonicity is necessary for the equilibrium being separating. In fact, suppose that $y(\eta)$ is not monotonous along the support of η. It can be easily proved that there are two different party *types*, η_1 and η_2, such that $y(\eta_1)=y(\eta_2)$. Clearly, this contradicts the equilibrium being separating.◇

Differentiability

Argument: differentiability is necessary for the equilibrium being separating.

By the inverse function theorem, the differentiability of a monotone function implies the differentiability of the inverse. Thus, in the remainder we prove that the inverse of $y(\eta)$ is differentiable. The proof is by contradiction. Thus, assume that there exist a point y_0 where where $y^{-1}(y)$ is not differentiable. Being y_0 optimal for type $y^{-1}(y_0)$, a small departure to the left of y_0 reduces the objective in Eq. (4):

$$\left\{\frac{\rho+\overline{z}}{\overline{z}}\left[1-y_0+y^{-1}(y_0)-T\right]-\frac{\rho\,(1-\lambda)}{\overline{z}}\left[-y_0+y^{-1}(y_0)-T\right]\widehat{\eta}'(y_0^-)\right\}dy\leq 0 \tag{A.2.6}$$

where $\widehat{\eta}'(y_0^-)$ stands for $\lim_{y\longrightarrow y_0^-}\widehat{\eta}'(y)$. Since $dy<0$, Inequality (A.2.6) requires

$$\frac{\rho+\overline{z}}{\overline{z}}\left[1-y_0+y^{-1}(y_0)-T\right]-\frac{\rho\,(1-\lambda)}{\overline{z}}\left[-y_0+y^{-1}(y_0)-T\right]\widehat{\eta}'(y_0^-)\geq 0 \tag{A.2.7}$$

Analogously, a small departure to the right of y_0 also reduces the objective in Eq. (4):

$$\left\{\frac{\rho+\overline{z}}{\overline{z}}\left[1-y_0+y^{-1}(y_0)-T\right]-\frac{\rho\,(1-\lambda)}{\overline{z}}\left[-y_0+y^{-1}(y_0)-T\right]\widehat{\eta}'(y_0^+)\right\}dy\leq 0 \tag{A.2.8}$$

where $\widehat{\eta}'(y_0^+)$ stands for $\lim_{y\longrightarrow y_0^+}\widehat{\eta}'(y)$. Since $dy>0$, Inequality (A.2.6) requires

$$\frac{\rho+\overline{z}}{\overline{z}}\left[1-y_0+y^{-1}(y_0)-T\right]-\frac{\rho\,(1-\lambda)}{\overline{z}}\left[-y_0+y^{-1}(y_0)-T\right]\widehat{\eta}'(y_0^+)\leq 0 \tag{A.2.9}$$

Suppose momentarily that $-y_0+y^{-1}(y_0)-T>0$ and recall that that $\widehat{\eta}(y)$ coincides with $y^{-1}(\eta)$. Inequalities (A.2.7) and (A.2.9) imply

$$\left(y^{-1}\right)'(y_0^-) \leq \frac{\rho+\overline{z}}{\rho(1-\lambda)}\frac{1-y_0+y^{-1}(y_0)-T}{-y_0+y^{-1}(y_0)-T} \leq \left(y^{-1}\right)'(y_0^+) \tag{A.2.10}$$

Since $y^{-1}(y)$ is not differentiable at y_0 either 1) $\left(y^{-1}\right)'(y_0^-) < \left(y^{-1}\right)'(y_0^+)$ or 2) $\left(y^{-1}\right)'(y_0^-) > \left(y^{-1}\right)'(y_0^+)$. Notice that Eq. (A.2.10) is only consistent with case 1). In this case, however, there is a range of *types* $y^{-1}(y_0)$ that satisfy Eq. (A.2.10). Yet, the fact that a range of types may set the same y_0 is not consistent with the equilibrium being separating.

Alternatively, one may suppose that $-y_0+y^{-1}(y_0)-T<0$. In this case, the inequalities in Eq. (A.2.10) are reversed but the argument used to rule out the jump in the derivative holds unchanged.⋄

A.3 Endogenous Information Acquisition (Eq. 17 and Figure 1)

In this section we solve Problem (15), which we report below for easy reference upon substituting the original variables $\sigma^2_{q^i_j}$ with $\lambda^i_j=\sigma^2_x/(\sigma^2_x+\sigma^2_{q^i_j})$, $j=A,B$:

$$\max_{\lambda^i_A,\lambda^i_B} E_{q^i_A,q^i_B,z}\left\{\max\left[E\left(U_A|y_A,q^i_A\right)+z, E\left(U_A|y_B,q^i_B\right)\right]\right\}-C(\lambda^i_A,\lambda^i_B) \qquad 0\leq\lambda^i_j\leq 1 \tag{A.3.1}$$

Where

$$C(\sigma^2_{q^i_A},\sigma^2_{q^i_B}) = -\frac{1}{2}\kappa\left[\log\left(1-\lambda^i_A\right)+\log\left(1-\lambda^i_B\right)\right] \tag{A.3.2}$$

$$U_j = y_j-\frac{1}{2}\left(-y_j+\eta_j+x_j-T\right)^2 \quad j=A,B \tag{A.3.3}$$

Compute U_j along the equilibrium policy $y(\eta_j)=1-T+\eta_j+D$ (D is exogenous to the individual, to save on notation we disregard the arguments of D). Then, compute the inner expectations in Equation (A.3.1):

$$E\left(U_j|y_j,q^i_j\right)$$

$$= y_j-\frac{1}{2}E\left[\left(-1-D+x_j\right)^2|y_j,q^i_j\right]$$

$$= y_j-\frac{1}{2}E\left[\left(1+D\right)^2-2\left(1+D\right)x_j+x_j^2|y_j,q^i_j\right]$$

$$= y_j-\frac{1}{2}\left(1+D\right)^2+\left(1+D\right)\lambda^i_j\left(q^i_j-\eta_j\right)-\frac{1}{2}\left[\sigma^2_x\left(1-\lambda^i_j\right)+\left(\lambda^i_j\right)^2\left(q^i_j-\eta_j\right)^2\right] \tag{A.3.4}$$

By using Eq.(A.3.4), one may state that

$$\max\left[E\left(U_A|y_A,q_A^i\right)+z,E\left(U_B|y_B,q_B^i\right)\right]=E\left(U_B^i|y_B,q_B^i\right)+\max\left[\Delta^i+z,0\right] \tag{A.3.5}$$

$$\Delta^i\equiv y_A-y_B+(1+D)\left[\lambda_A^i\left(q_A^i-\eta_A\right)-\lambda_B^i\left(q_B^i-\eta_B\right)\right]+$$

$$-\tfrac{1}{2}\left[\sigma_x^2\left(1-\lambda_A^i\right)+\left(\lambda_A^i\right)^2\left(q_A^i-\eta_A\right)^2-\sigma_x^2\left(1-\lambda_B^i\right)-\left(\lambda_B^i\right)^2\left(q_B^i-\eta_B\right)^2\right]$$

Further, by using Eq.(A.3.4), the outer expectation of $E\left(U_B^i|y_B,q_B^i\right)$ is

$$E_{q_A^i,q_B^i,z}\left[E\left(U_B|y_B,q_B^i\right)\right]$$

$$=y_B-\frac{1}{2}(1+D)^2-\frac{1}{2}\sigma_x^2\left(1-\lambda_B^i\right)-\frac{1}{2}\left(\lambda_B^i\right)^2E_{q_A^i,q_B^i,z}\left[\left(q_B^i-\eta_B\right)^2\right]$$

$$=y_B-\frac{1}{2}(1+D)^2-\frac{1}{2}\sigma_x^2\left(1-\lambda_B^i\right)-\frac{1}{2}\left(\lambda_B^i\right)^2\left(\sigma_x^2+\sigma_{q_B^i}^2\right)$$

$$=y_B-\frac{1}{2}(1+D)^2-\frac{1}{2}\sigma_x^2\left(1-\lambda_B^i\right)-\frac{1}{2}\sigma_x^2\lambda_B^i$$

$$=y_B-\frac{1}{2}(1+D)^2-\frac{1}{2}\sigma_x^2 \tag{A.3.6}$$

Eq. (A.3.6) proves that $E_{q_A^i,q_B^i,z}\left[E\left(U_B|y_B,q_B^i\right)\right]$ is independent from $(\sigma_{q_A^i}^2,\sigma_{q_B^i}^2)$. Hence, On the basis of Eq.(A.3.5), the Problem (A.3.1) can be expressed in the following compact form:

$$\max_{\sigma_{q_A^i}^2,\sigma_{q_B^i}^2}E_{q_A^i,q_B^i,z}\left\{\max\left[\Delta^i+z,0\right]\right\}-C(\lambda_A^i,\lambda_B^i) \tag{A.3.7}$$

Let the support of z be sufficiently large, the expectation in Problem (A.3.7) can be expressed as follows:

$$E_{q_A^i,q_B^i,z}\left\{\max\left[\Delta^i+z,0\right]\right\}=E_{q_A^i,q_B^i}\left\{\int_{-\Delta^i}^{\frac{\overline{z}}{2}}\frac{1}{\overline{z}}\left(\Delta^i+z\right)dz\right\}=\frac{1}{\overline{z}}E_{q_A^i,q_B^i}\left\{\frac{\overline{z}^2}{8}+\frac{1}{2}\left(\Delta^i\right)^2+\frac{\overline{z}}{2}\Delta^i\right\} \tag{A.3.8}$$

Since $E_{q_A^i,q_B^i}\left(\Delta^i\right)=y_A-y_B$, the only term on the RHS of Eq. (A.3.8) that depends on $(\sigma_{q_A^i}^2,\sigma_{q_B^i}^2)$ is $E_{q_A^i,q_B^i}\left[\left(\Delta^i\right)^2\right]$. Hence, the Problem (A.3.7) can be expressed as follows:

$$\max_{\sigma^2_{q^i_A}, \sigma^2_{q^i_B}} E_{q^i_A, q^i_B, z} \frac{1}{\overline{z}} \frac{1}{2} \left[\left(\Delta^i \right)^2 \right] - C(\lambda^i_A, \lambda^i_B) \tag{A.3.9}$$

After some algebra, we obtain

$$E_{q^i_A, q^i_B, z} \left[\left(\Delta^i \right)^2 \right] = (y_A - y_B)^2 + (1+D)^2 \sigma^2_x \lambda^i_A + \frac{1}{2} \sigma^4_x \left(\lambda^i_A \right)^2 + (1+D)^2 \sigma^2_x \lambda^i_B + \frac{1}{2} \sigma^4_x \left(\lambda^i_B \right)^2 \tag{A.3.10}$$

Substitute Equation (A.3.10) in Problem (A.3.9) and observe that, for symmetry, the solution for λ^i_A must be equal to the solution for λ^i_B. Let us denote such solution with λ^i. The necessary condition for an optimum inside the interval $(0, 1)$ is

$$(1+D)^2 \sigma^2_x + \sigma^4_x \lambda^i = \kappa \overline{z} \frac{1}{1 - \lambda^i} \tag{A.3.11}$$

This equation is reported in the main text as Equation (17).

Equation (A.3.11) is a second order equation with respect to λ^i. Let λ^i_{lower} and λ^i_{upper} be the lower and upper solutions of this equations, it turns out that only λ^i_{upper} satisfies the second order sufficient condition. Further, recall that λ^i_{lower} and λ^i_{upper} are given by the standard formula

$$\lambda^i = \frac{-\left[(1+D)^2 \sigma^2_x - \sigma^4_x \right] \pm \sqrt{\left[(1+D)^2 \sigma^2_x - \sigma^4_x \right]^2 - 4\sigma^4_x \left[\kappa \overline{z} - (1+D)^2 \sigma^2_x \right]}}{2\sigma^4_x} \tag{A.3.12}$$

In the parametrization used for the plots in Figure 1, we have set $\sigma^2_x < 1$. Thus, the General Solution (A.3.12) implies that $\lambda^i_{lower} < 0$ while λ^i_{upper} is positive only if $(1+D)^2 \sigma^2_x > \kappa \overline{z}$, that is only if the cost parameter κ is sufficiently small. Notice that this condition holds along the range of values assigned to κ in the graph. Further, notice that $\lambda^i_{upper} < 1$ holds for all parameter values. Thus, consistent with the Equation (A.3.11) representing the optimality condition for an internal optimum, in computations used to depict Figure 1 it holds $\lambda^i_{upper} \in (0, 1)$.

Table A1: Variables and Descriptive Statistics

Variables	Description (Source)	Mean	Std. Dev.	Min	Max
PRIMARY BALANCE	Government primary balance (excluding net interest payments) as a percentage of GDP (OECD Economic Outlook).	1.673	3.672	-13.528	16.102
TOTAL OUTLAYS	Government total outlays as a percentage of GDP(OECD Economic Outlook).	-0.039	4.379	-13.507	18.787
TOTAL TAX AND NO-TAX RECEIPTS	Government tax and no-tax receipts as a percentage of GDP (OECD Economic Outlook).	43.644	6.521	31.219	58.112
LITERACY	Indicator (0-10 scale) based on experts interviews on population economic literacy (World Competitiveness Yearbook).	43.606	7.399	30.121	58.898
POLARIZATION	Dummy that equals one if there is hegemony in the cabinet composition (Elaborations on Comparative Political Dataset).	5.820	1.227	3.010	8.160
DEBT	Gross government debt (financial liabilities) as a percentage of GDP (OECD Economic Outlook).	0.661	0.474	0.000	1.000
INTEREST	Long-term interest rate on government bonds (OECD Economic Outlook).	65.272	33.032	11.280	171.128
UNEMPLOYMENT	Unemployment rate as a percentage of civilian labour force (OECD Employment and Labour Market Statistics).	4.651	1.435	1.003	11.195
GDP GROWTH	Annual percentage growth rate of GDP per capita based on constant local currency (OECD Main Economic Indicators).	5.964	2.493	2.007	15.692
INFLATION	Growth of consumer price index (CPI), all items, as a percentage change from previous year (OECD Main Economic Indicators).	1.971	1.760	-4.507	9.484
OPEN	Total trade (sum of import and export) as a percentage of GDP, in current prices (Penn World Table 8.0).	2.359	1.403	-0.900	12.655
POPULATION	Total population (millions) (OECD Employment and Labour Market Statistics).	85.576	54.435	18.756	319.554
EDUCATION	Average number of years of education of women and men aged 25 and older (Gakidou et al.,2010).	37.842	62.650	0.277	304.094
CICLICALLY ADJ. PRIM. BALANCE	Cyclically adjusted government primary balance (excluding net interest payments) as a percentage of potential GDP (OECD Economic Outlook).	22.518	3.269	12.500	28.200

Table A2: Cyclically-adjusted primary balance, government polarization and economic literacy

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	*CYCLICALLY-ADJUSTED PRIMARY BALANCE*				
$LITERACY_{t-1}$	-0.68	-0.63	-0.70	-0.68	0.10
	(0.76)	(0.68)	(0.73)	(0.73)	(0.47)
$POLARIZATION_{t-1}$	-6.86*	-6.86*	-6.48*	-6.28*	-2.91***
	(3.44)	(3.14)	(3.18)	(3.12)	(0.92)
$(POLAR. * LITERACY)_{t-1}$	1.36*	1.38*	1.32*	1.28*	0.58***
	(0.70)	(0.64)	(0.66)	(0.65)	(0.19)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	207	207	207	207	198
$R - squared$	0.346	0.372	0.398	0.414	0.432
Number of countries	23	23	23	23	22

Notes:The table reports FE regression coefficients and country-level clustered robust standard errors (in brackets).Time and country dummies are included in the estimates (coefficients are omitted in the table). All regressions are estimated with an intercept term.Controls are the same used in Table 1. The values of EDUCATION are not available for Iceland (column (5)). *p $<$ 0.10; **p $<$ 0.05; ***p $<$ 0.01.

Table A3: test on unobserved heterogeneity

	(1)	(2)
Dep. Var.	Predicted Primary Balance	Primary Balance
$LITERACY_{t-1}$	-0.23	-0.04
	(0.14)	(0.68)
$POLARIZATION_{t-1}$	-0.47	-5.17*
	(0.58)	(2.95)
$(POLARIZATION * LITERACY)_{t-1}$	0.07	1.06*
	(0.09)	(0.59)
Predicted Pr. Bal.		0.91***
		(0.31)
Controls	Yes	Yes
Observations	207	207
$R-squared$	0.88	0.46
Number of countries	23	23

Notes:The table reports FE regression coefficients and country-level clustered robust standard errors (in brackets). The dependent variables are the predicted primary balance over GDP (first column) and the primary balance over GDP (last column). Time and country dummies are included in the estimates (coefficients are omitted in the table). All regressions are estimated with an intercept term. Controls are the same used in Table 1. *p $<$ 0.10; **p $<$ 0.05; ***p $<$ 0.01.

We implement two tests proposed by Chetty et al. (2011) that are more suitable for our research setting. First, for each year of our sample we estimated OLS regressions where the dependent variable is the primary budget, and the covariate is the above initially omitted variable. Then, we estimate our main models by substituting the dependent variable (PRIMARY BALANCE) with its predicted value (Predicted PRIMARY BALANCE).

In column (1) of Table A3, we find that our main results do not hold, and thus our baseline results do not seem to be driven by unobserved correlation between fiscal budget and government quality. In column (2) of Table A3, we include the Predicted PRIMARY BALANCE as additional regressor in our main models, and we find that: i) this predicted balance is strongly significant, meaning that government quality is a strong predictor of the fiscal budget; but ii) our main results hold, and thus the degree of bias due to potential unobserved heterogeneity in our baseline models is likely to be small.

Table A4: Falsification test

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	*PRIMARY BALANCE*				
$LITERACY_{t-1}$	-0.26	-0.28	-0.22	-0.18	-0.09
	(0.43)	(0.41)	(0.41)	(0.39)	(0.40)
$POLARIZATION_{t-1}$	-0.52	-1.04	-1.14	-0.93	-0.01
	(1.47)	(1.50)	(1.38)	(1.39)	(1.29)
$(POLAR. * LITERACY)_{t-1}$	0.11	0.21	0.23	0.22	0.04
	(0.30)	(0.30)	(0.28)	(0.28)	(0.26)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	207	207	207	207	198
$R-squared$	0.302	0.326	0.412	0.450	0.451
Number of countries	23	23	23	23	22

Notes:The table reports FE regression coefficients and country-level clustered robust standard errors (in brackets). Time and country dummies are included in the estimates (coefficients are omitted in the table). All regressions are estimated with an intercept term. Controls are the same used in Table 1. The values of EDUCATION are not available for Iceland (column (5)). *p $<$ 0.10; **p $<$ 0.05; ***p $<$ 0.01.

We use three falsification tests and randomize the variable LITERACY across: i) pooled observations in the dataset; ii) years within countries; and iii) countries within years. As expected, our main results vanish. This seems to exclude that the effect we found is spurious. Results of the more general test i) are reported in Table A4.