Online Appendix for

# Do Elections Moderate or Polarize Political Rhetoric?

Tito Boeri* Nina Nikiforova† Guido Tabellini‡

July 27, 2026

*Department of Economics and IGIER, Bocconi University; IZA; CEP-LSE
†Department of Economics, Bocconi University
‡Department of Economics and IGIER, Bocconi University; CEPR; CES-Ifo

# A Proof of Proposition 1

*Proof.* Candidate $A$ chooses $(\mathbf{x}^A, \mathbf{y}^A, s^A)$ to maximize equation (6), taking $B$'s communication as given. A fraction $\alpha$ of voters of each type is reached by the communication of both candidates, a fraction $1-\alpha$ of $a$-voters is reached only by $A$, and a fraction $1-\alpha$ of $b$-voters only by $B$; voters not reached by a candidate's communication behave as if it were zero. Writing $\Delta^{v,r}$ for the advantage $\Delta^v = V(\mathbf{g}^v, \mathbf{g}^A) - V(\mathbf{g}^v, \mathbf{g}^B)$ of a type-$v$ voter with reach status $r \in \{rr, rA, rB\}$, the vote share therefore decomposes by audience as

$$\pi^A = \frac{1}{2} + \frac{\phi}{2}\Big\{\alpha \underbrace{\left[\Delta^{a,rr} + \Delta^{b,rr}\right]}_{\substack{\text{voters reached by}\\ \text{both candidates}}} + (1-\alpha) \underbrace{\Delta^{a,rA}}_{\substack{A\text{'s exclusive}\\ \text{audience}}} + (1-\alpha) \underbrace{\Delta^{b,rB}}_{\substack{B\text{'s exclusive}\\ \text{audience}}} \Big\}. \tag{A1}$$

Each audience term is an explicit function of the communication efforts. Since voter utility is quadratic, $V(\mathbf{g}^v, \mathbf{g}^P) = -\frac{1}{2}\sum_k \sigma_k^v (g_k^v - g_k^P)^2$, the difference of squares gives

$$\Delta^{v,r} = \sum_k \sigma_k^{v,r}\, \delta_k^{v,r}, \qquad \delta_k^{v,r} \equiv \frac{1}{2}\left(g_k^A - g_k^B\right)\left(2g_k^v - g_k^A - g_k^B\right), \tag{A2}$$

where all perceptions on the right are evaluated at the values held by audience $r$ : voters reached by both candidates hold the perceptions of the communication technology in the text, while voters in an exclusive audience hold the same perceptions with the unreached candidate's communication set to zero. Explicitly, for $k = 1, 2$ and with $\sigma_2^v = 1 - \sigma_1^v$ throughout:

$$\begin{aligned}
rr\ (v = a, b):& \quad g_k^A = \hat{g}_k^A + x_k^A,\ g_k^B = \hat{g}_k^B + x_k^B,\ g_k^v = \hat{g}_k^v + \lambda^v(y_k^A + y_k^B),\ \sigma_1^v = \bar{\sigma}^v + s^A + s^B;\\
rA\ (v = a):& \quad g_k^A = \hat{g}_k^A + x_k^A,\quad g_k^B = \hat{g}_k^B,\quad g_k^a = \hat{g}_k^a + y_k^A,\quad \sigma_1^a = \bar{\sigma}^a + s^A;\\
rB\ (v = b):& \quad g_k^A = \hat{g}_k^A,\quad g_k^B = \hat{g}_k^B + x_k^B,\quad g_k^b = \hat{g}_k^b - y_k^B,\quad \sigma_1^b = \bar{\sigma}^b + s^B.
\end{aligned}$$

Thus candidate $A$'s instruments enter $\Delta^{a,rr}, \Delta^{b,rr}$ (all three channels) and $\Delta^{a,rA}$ (own position $x_k^A$, bliss points $y_k^A$, and salience $s^A$, with $B$ perceived at its primitive positions), while $\Delta^{b,rB}$ contains none of $A$'s instruments and therefore drops from all of $A$'s first-order conditions.

We first derive the five FOCs at the $\mathcal{T}$-symmetric profile that the equilibrium must satisfy; we then verify each claim of the proposition. The FOCs are obtained by differentiating (A1), using (A2), instrument by instrument.

**Setup.** The model is invariant under the joint transformation $\mathcal{T} : A \leftrightarrow B,\ a \leftrightarrow b,\ k \leftrightarrow k' = \{1, 2\} \setminus \{k\}$ (the dimension swap is required because $\bar{\sigma}^a + \bar{\sigma}^b = 1$ preserves the salience pattern when voter types and dimensions are swapped together). Under $\mathcal{T}$, the perceived-position instrument flips sign (since $\hat{g}^A = -\hat{g}^B$), the bliss-point instrument does not (since $y_k^A + y_k^B$ enters symmetrically), and the salience instrument flips sign (since $\sigma_1$ maps to $1-\sigma_1$ under the dimension swap). By uniqueness, the transformed equilibrium must coincide with the original one. Hence the equilibrium satisfies

$$x_k^A = -x_{k'}^B, \quad y_k^A = y_{k'}^B, \quad s^A = -s^B. \tag{A3}$$

We write $u \equiv x_1^A$, $w \equiv x_2^A$, $\mu \equiv y_1^A$, $\nu \equiv y_2^A$, and $\tau \equiv \mu + \nu$. The salience weights at the $\mathcal{T}$-symmetric profile are: in the reached-by-both subgroup (rr), $\sigma_1^v = \bar{\sigma}^v$; in the $a$-only subgroup (rA), $\sigma_1^a = \frac{1}{2} + \epsilon + s^A$; in the $b$-only subgroup (rB), $\sigma_1^b = \frac{1}{2} - \epsilon - s^A$.

Direct computation, using $\sum_v \sigma_k^v \hat{g}_k^v = \pm 2\epsilon(\varpi + \beta)$ (positive on dim 1, negative on dim 2) and $\sum_v \sigma_k^v \lambda^v = \pm 2\epsilon$, yields the five marginal benefits:

$$\begin{aligned}
\frac{\partial \pi^A}{\partial u} &= \alpha\frac{\phi}{2}\left[2\epsilon(\varpi+\beta+\tau) - (\varpi+u)\right] + (1-\alpha)\frac{\phi}{2}\sigma_1^a(\beta+\mu-u), && \text{(A4)}\\
\frac{\partial \pi^A}{\partial w} &= \alpha\frac{\phi}{2}\left[-2\epsilon(\varpi+\beta+\tau) - (\varpi+w)\right] + (1-\alpha)\frac{\phi}{2}(1-\sigma_1^a)(\beta+\nu-w), && \text{(A5)}\\
\frac{\partial \pi^A}{\partial \mu} &= \alpha\phi\epsilon(2\varpi+u+w) + (1-\alpha)\frac{\phi}{2}\sigma_1^a(2\varpi+u), && \text{(A6)}\\
\frac{\partial \pi^A}{\partial \nu} &= -\alpha\phi\epsilon(2\varpi+u+w) + (1-\alpha)\frac{\phi}{2}(1-\sigma_1^a)(2\varpi+w), && \text{(A7)}\\
\frac{\partial \pi^A}{\partial s^A} &= -\alpha\phi(2\varpi+u+w)(u-w) + (1-\alpha)\frac{\phi}{2}\left[\delta_1^{a,rA} - \delta_2^{a,rA}\right], && \text{(A8)}
\end{aligned}$$

where $\delta_k^{a,rA} = \frac{1}{2}(2\varpi + x_k^A)(2\varpi + 2\beta + 2y_k^A - x_k^A)$, using $g_1^A - g_1^B = g_2^A - g_2^B = 2\varpi + u + w$ at the $\mathcal{T}$-symmetric profile, and the rr-term in (A8) uses $\sum_v(\delta_1^v - \delta_2^v)|_{rr} = -2(2\varpi+u+w)(u-w)$, obtained by direct expansion of $\delta_k^v = \frac{1}{2}(g_k^A - g_k^B)(2g_k^v - g_k^A - g_k^B)$ with $g_1^A + g_1^B = u - w$ and $g_2^A + g_2^B = -(u-w)$. The equilibrium satisfies $C'(z_i) = \gamma \cdot \partial\pi^A/\partial z_i$ for each $z_i \in \{u, w, \mu, \nu, s^A\}$. Since $C'$ is odd, strictly increasing, and equals 0 at the origin, $z_i$ has the same sign as the corresponding marginal benefit.

**Parts (i) and (ii):** $\epsilon = 0$. At $\epsilon = 0$ the model is additionally invariant under the dimension-swap $\mathcal{S} : k \leftrightarrow k'$ alone, which under (3) sends $s^A$ to $-s^A$. At an equilibrium fixed by $\mathcal{S}$, therefore, $u = w$, $\mu = \nu$, and $s^A = 0$; moreover $s^B = -s^A = 0$ by (A3). Substituting $\sigma_1^a = \frac{1}{2}$ and $u = w$, $\mu = \nu$ into (A8) confirms that both the rr- and rA-terms vanish at this point, so the salience FOC is satisfied identically. The remaining FOCs collapse to a two-variable system for $(u,\mu)$ :

$$C'(u)=\gamma\frac{\phi}{2}\left[-\alpha(\varpi+u)+\frac{1-\alpha}{2}(\beta+\mu-u)\right],\qquad C'(\mu)=\gamma\frac{\phi}{4}(1-\alpha)(2\varpi+u).\qquad \text{(A9)}$$

*Part (i):* $\alpha=1$. Only the rr-term survives. The bliss-point FOC gives $C'(\mu)=0$, hence $\mu=0$ and $y_k^P=0$. The position FOC becomes $C'(u)=-\gamma(\phi/2)(\varpi+u)$, whose right-hand side is strictly negative at $u=0$, zero at $u=-\varpi$, and strictly decreasing in $u$; together with $C'$ strictly increasing, this gives a unique solution $u\in(-\varpi,0)$, so $x_k^A=-x_k^B=u<0$. Implicit differentiation of $C'(u)+\gamma(\phi/2)(\varpi+u)=0$ yields $\partial u/\partial\gamma=-\frac{\phi}{2}(\varpi+u)/[C''(u)+\gamma\phi/2]<0$ (using $\varpi+u>0$), so $\partial|u|/\partial\gamma>0$, confirming (iv) in this regime.

*Part (ii):* $\alpha<1$. The bliss-point FOC $C'(\mu)=\gamma(1-\alpha)(\phi/4)(2\varpi+u)$ determines $\mu$ as a strictly increasing function $\mu=h(u)$ of $u$ (since $C'$ is strictly increasing, hence invertible, and the right-hand side is increasing in $u$), with $h(0)>0$ because $C'(h(0))=\gamma(1-\alpha)\phi\varpi/2>0=C'(0)$. Substituting $\mu=h(u)$ into the position FOC yields a single equation in $u$ :

$$F(u)\equiv C'(u)-\gamma\frac{\phi}{2}G_u(u,h(u))=0,\qquad G_u(u,\mu)\equiv-\alpha(\varpi+u)+\frac{1-\alpha}{2}(\beta+\mu-u).$$

Implicit differentiation of $C'(h(u))=\gamma(1-\alpha)(\phi/4)(2\varpi+u)$ gives $C''(h(u))h'(u)=\gamma(1-\alpha)\phi/4$, so

$$F'(u)=C''(u)-\gamma\frac{\phi}{2}\left[\frac{\partial G_u}{\partial u}+\frac{\partial G_u}{\partial\mu}h'(u)\right]=\frac{\det J(u)}{C''(h(u))},$$

where $J(u)\equiv\operatorname{diag}\left(C''(u),C''(h(u))\right)-\gamma\frac{\phi}{2}\nabla G(u,h(u))$ is the same $2\times2$ Jacobian used below for the comparative statics, here evaluated along the entire curve $\mu=h(u)$ rather than only at the equilibrium. Under the maintained convexity assumption, taken — as is standard for this type of argument — to hold along this curve rather than only at the equilibrium point, $\det J(u)>0$ everywhere, so $F$ is strictly increasing and has at most one root; by the maintained existence of an interior equilibrium, exactly one.

To locate this root, evaluate $F$ at $u=0$. Since $C'(0)=0$ and $h(0)>0$,

$$F(0)=-\gamma\frac{\phi}{2}G_u(0,h(0)),$$

where

$$G_u(0,h(0))=-\alpha\varpi+\frac{1-\alpha}{2}\left(\beta+h(0)\right)>-\alpha\varpi+\frac{1-\alpha}{2}\beta=\frac{1}{2}\left[(1-\alpha)\beta-2\alpha\varpi\right],$$

For $\alpha<\bar{\alpha}\equiv\beta/(2\varpi+\beta)$, we have $G_u(0,h(0))>0$, so $F(0)<0$; since $F$ is strictly increasing, its unique root satisfies $u^* > 0$. Since $h$ is increasing and $u^* > 0$, $\mu^* = h(u^*) > h(0) > 0$. Hence $x_k^A = -x_k^B = u^* > 0$, $y_k^P = \mu^* > 0$; set $\hat{\alpha} = \bar{\alpha}$ at $\epsilon = 0$.

For the comparative statics of part (iv) discussed below, system (A9) defines $(u, \mu)$ implicitly as functions of $\gamma$. At $\alpha < \bar{\alpha}$ the right-hand sides $G_u$ and $G_\mu = \frac{1-\alpha}{2}(2\varpi + u)$ are strictly positive at the equilibrium. The Jacobian $J$ introduced above, now evaluated at the equilibrium, is a Z-matrix (both off-diagonal entries equal $-\gamma\phi(1-\alpha)/4 < 0$) with dominant diagonal under the maintained convexity assumption, hence an M-matrix with non-negative inverse; applied to the positive vector $\frac{\phi}{2}(G_u, G_\mu)$ this yields $\partial u/\partial\gamma > 0$ and $\partial\mu/\partial\gamma > 0$, confirming (iv) in this regime.

**Part (iii): $\alpha < \hat{\alpha}$ and $\epsilon > 0$ sufficiently small.** The $\mathcal{S}$-symmetry breaks because $\bar{\sigma}^a \neq 1/2$. We work with the full five FOCs (A4)–(A8).

*Step 1: first-order expansion in $\epsilon$.* Under SOC the equilibrium is smooth in $\epsilon$ (implicit function theorem). Since the model with $-\epsilon$ is the $\mathcal{S}$-image of the model with $\epsilon$, the differences $\Delta x \equiv u - w$ and $\Delta y \equiv \mu - \nu$ and the salience effort $s^A$ are odd functions of $\epsilon$, hence $O(\epsilon)$, while $u + w$ and $\mu + \nu$ are even, so $u, w = u_0 + O(\epsilon)$ and $\mu, \nu = \mu_0 + O(\epsilon)$ with $(u_0, \mu_0)$ the $\epsilon = 0$ equilibrium of part (ii). Linearizing the differences (A4)−(A5) and (A6)−(A7), and equation (A8), around $(u_0, \mu_0)$ — using $\sigma_1^a = \frac{1}{2} + \epsilon + s^A$, $\tau = 2\mu_0 + O(\epsilon^2)$, and $\delta_1^{a,rA} - \delta_2^{a,rA} = (\beta + \mu_0 - u_0)\Delta x + (2\varpi + u_0)\Delta y + o(\epsilon)$ — yields, up to $o(\epsilon)$ terms,

$$\begin{aligned}\left[C''(u_0) + \gamma\tfrac{\phi(1+\alpha)}{4}\right]\Delta x - \gamma\tfrac{(1-\alpha)\phi}{4}\Delta y - \gamma(1-\alpha)\phi(\beta + \mu_0 - u_0)\, s^A \;&=\\ \gamma\epsilon\phi\left[2\alpha(\varpi + \beta + 2\mu_0) + (1-\alpha)(\beta + \mu_0 - u_0)\right]&,\end{aligned}$$

$$\begin{aligned}-\gamma\tfrac{(1-\alpha)\phi}{4}\Delta x + C''(\mu_0)\Delta y - \gamma(1-\alpha)\phi(2\varpi + u_0)\, s^A \;&=\\ \gamma\epsilon\phi\left[4\alpha(\varpi + u_0) + (1-\alpha)(2\varpi + u_0)\right]&,\end{aligned}$$

$$-\gamma\kappa_u\Delta x - \gamma\kappa_\mu\Delta y + C''(0)\, s^A \;=\; 0,$$

where $\kappa_u \equiv \frac{(1-\alpha)\phi}{2}(\beta + \mu_0 - u_0) - 2\alpha\phi(\varpi + u_0)$ and $\kappa_\mu \equiv \frac{(1-\alpha)\phi}{2}(2\varpi + u_0) > 0$. Both right-hand sides in the first two rows are strictly positive (recall $\beta + \mu_0 - u_0 > 0$, since $C'(u_0) > 0$ in (A9)). If $\kappa_u > 0$ — which holds for $\alpha$ sufficiently small, since $\kappa_u \to \frac{\phi}{2}(\beta + \mu_0 - u_0) > 0$ as $\alpha \to 0$ — the coefficient matrix is a Z-matrix with dominant diagonal under the maintained convexity assumption, hence an M-matrix with non-negative inverse. Therefore $\Delta x > 0$ and

$\Delta y > 0$, both of order $\epsilon$, and the third row gives

$$s^A = \frac{\gamma}{C''(0)} [\kappa_u \Delta x + \kappa_\mu \Delta y] + o(\epsilon) > 0. \tag{A10}$$

By (A3) these deliver the orderings $x_1^A = -x_2^B > x_2^A = -x_1^B$, $y_1^A = y_2^B > y_2^A = y_1^B$, and $s^A = -s^B > 0$.

*Step 2: positivity of* $u, w, \mu, \nu$. At $\epsilon = 0$ and $\alpha < \bar{\alpha} = \beta/(2\varpi + \beta)$, part (ii) establishes $u = w > 0$ and $\mu = \nu > 0$. The equilibrium depends continuously on $\epsilon$ under SOC, so each of $u, w, \mu, \nu$ remains strictly positive in a neighborhood of $\epsilon = 0$.

Define

$$\hat{\alpha}(\epsilon) \equiv \sup \left\{ a \in (0, \bar{\alpha}] : u, w, \mu, \nu, s^A > 0 \text{ at every } \alpha \in (0, a) \right\},$$

the largest initial interval of $\alpha$ on which all five inequalities hold, so that the conclusions of part (iii) hold for every $\alpha < \hat{\alpha}(\epsilon)$ by construction. The preceding steps show that this set is nonempty, so $\hat{\alpha}(\epsilon) > 0$ : $\kappa_u$ is continuous in $\alpha$ with $\kappa_u \to \frac{\phi}{2}(\beta + \mu_0 - u_0) > 0$ as $\alpha \to 0$, so $\kappa_u > 0$ on a compact interval $[0, a_0]$ with $0 < a_0 < \bar{\alpha}$; since $(u_0, \mu_0)$ and hence the M-matrix bounds of Step 1 and the continuity bounds of Step 2 vary continuously with $\alpha$, the $o(\epsilon)$ remainders are uniform in $\alpha$ over $[0, a_0]$, and all five inequalities hold at every $\alpha \in (0, a_0)$ once $\epsilon$ is sufficiently small. Hence $\hat{\alpha}(\epsilon) \geq a_0$. Note that the requirement $s^A > 0$ binds through the sign of $\kappa_u \Delta x + \kappa_\mu \Delta y$ in (A10), which is independent of $\epsilon$ at leading order; hence, unlike the thresholds governing $u, w, \mu, \nu > 0$, $\hat{\alpha}(\epsilon)$ need not converge to $\bar{\alpha}$ as $\epsilon \to 0$.

**Part (iv): comparative statics in $\gamma$.** The equilibrium is defined implicitly by the five-equation system

$$C'(z_i) = \gamma \cdot G_i(z; \alpha, \epsilon), \quad i \in \{u, w, \mu, \nu, s^A\}, \tag{A11}$$

where $G_i$ is the corresponding right-hand side in (A4)–(A8). Under SOC the Jacobian $J \equiv \text{diag}(C'') - \gamma \nabla_z G$ is invertible at the equilibrium, and the implicit function theorem yields

$$\frac{\partial z}{\partial \gamma} = J^{-1} G(z).$$

We verify the claim regime by regime.

*Regimes (i) and (ii)* ($\epsilon = 0$). Since the model is $\mathcal{S}$-invariant at every $\gamma$, the equilibrium path $\gamma \mapsto z(\gamma)$ stays on the symmetric manifold $u = w$, $\mu = \nu$, $s^A = 0$, and the comparative statics reduce to those of the lower-dimensional systems analyzed above: part (i) gives $\partial |u|/\partial \gamma > 0$ by direct implicit differentiation, and part (ii) gives $\partial u/\partial \gamma, \partial \mu/\partial \gamma > 0$ from the two-variable M-matrix argument. Instruments equal to zero remain zero. This confirms (iv) at $\epsilon = 0$.

*Regime (iii)* ($\epsilon > 0$ small). For the position and bliss-point instruments, $\partial z_i/\partial\gamma$ is continuous in $(\gamma, \epsilon)$ under SOC and strictly positive at $\epsilon = 0$ by the previous paragraph; hence, on any compact set of $\gamma$, for $\epsilon$ sufficiently small $\partial u/\partial\gamma, \partial w/\partial\gamma, \partial\mu/\partial\gamma, \partial\nu/\partial\gamma > 0$, and since these four instruments are strictly positive by part (iii), $\partial|z_i|/\partial\gamma > 0$ for each of them. For the salience instrument continuity is uninformative, because $\partial s^A/\partial\gamma \to 0$ as $\epsilon \to 0$. By (A10), $s^A = \epsilon\,\gamma\Psi(\gamma)/C''(0) + o(\epsilon)$, where $\epsilon\Psi(\gamma) \equiv \kappa_u\Delta x + \kappa_\mu\Delta y > 0$ is evaluated at the $\epsilon = 0$ equilibrium associated with $\gamma$. Hence $\partial|s^A|/\partial\gamma > 0$ for $\epsilon$ small if and only if $\gamma\Psi(\gamma)$ is strictly increasing in $\gamma$. This is a joint restriction on the cost function and the remaining parameters of the model; Lemma 1 below establishes it for quadratic costs and $\alpha$ sufficiently small. □

**Lemma 1.** *Let $C(z) = \frac{c}{2}z^2$ and $\bar\alpha \equiv \beta/(2\varpi + \beta)$. There exists $\breve\alpha \in (0, \bar\alpha]$ such that for every $\alpha \in [0, \breve\alpha)$, at every interior equilibrium satisfying the maintained conditions (positivity of the leading principal minors of the symmetric and antisymmetric Jacobians, $\kappa_u > 0$, $\beta > 0$, and $\sigma_1^a, \sigma_1^b \in (0, 1)$), $\gamma\Psi(\gamma)$ is strictly increasing in $\gamma$; hence $\partial|s^A|/\partial\gamma > 0$ for $\epsilon$ sufficiently small.*

*Proof. Step 1: decoupling at $\alpha = 0$.* At $\alpha = 0$ only the rA and rB groups matter, so the two dimensions interact only through the salience weight $\sigma \equiv \sigma_1^a$. Given $\sigma$, dimension 1's equilibrium $(u, \mu)$ solves the linear system $u = \frac{\eta}{2}(\beta + \mu - u)$, $\mu = \frac{\eta}{2}(2 + u)$ with $\eta = \gamma\sigma$, and dimension 2 solves the same system with $\eta = \gamma(1 - \sigma)$. Explicitly,

$$u^*(\eta) = \frac{(\eta/2)(\beta + \eta)}{D(\eta)}, \qquad \mu^*(\eta) = \eta + \frac{\eta}{2}u^*(\eta), \qquad D(\eta) = 1 + \frac{\eta}{2} - \frac{\eta^2}{4},$$

valid on the per-dimension SOC region $D(\eta) > 0$, i.e. $\eta < \bar\eta \equiv 1 + \sqrt{5}$.

*Step 2: the salience response as a scalar fixed point.* The salience FOC becomes $s = \frac{\gamma}{2}\left[\Phi(\gamma\sigma) - \Phi(\gamma(1 - \sigma))\right]$, where

$$\Phi(\eta) \equiv \frac{1}{2}\big(2 + u^*(\eta)\big)\big(2 + 2\beta + 2\mu^*(\eta) - u^*(\eta)\big)$$

is the advantage term of a dimension with weight parameter $\eta$, evaluated along the curve $(u^*, \mu^*)$. Since $\sigma = \frac{1}{2} + \epsilon + s$, linearizing at $\epsilon = 0$ gives $s = N(\gamma)(\epsilon + s) + o(\epsilon)$ with

$$N(\gamma) \equiv \gamma^2\Phi'(\gamma/2),$$

so that

$$\frac{s}{\epsilon} = \frac{N(\gamma)}{1 - N(\gamma)} + o(1),$$

and the salience second-order condition is exactly $N(\gamma) < 1$. Direct computation gives

$$\Phi'(\eta) = \frac{-4\left[\beta^2(\eta^3 + 6\eta^2 + 8) + 8\beta\eta(\eta^2 + 3\eta + 6) + 4(3\eta^3 + 12\eta^2 + 12\eta + 8)\right]}{(\eta^2 - 2\eta - 4)^3} > 0$$

on the SOC region, since there $\eta^2 - 2\eta - 4 = -4D(\eta) < 0$ and the bracket has all positive coefficients; hence $N > 0$ and $s > 0$, consistent with part (iii). As $\eta \to \bar{\eta}^-$, $\Phi'(\eta) \to +\infty$, so $N$ reaches 1 strictly before the position-block minor $D$ vanishes: the binding constraint defining the upper endpoint $\bar{\gamma}(0)$ of the admissible $\gamma$-interval is the salience minor $1 - N$, and the position-block minors have strict slack near $\bar{\gamma}(0)$.

*Step 3: monotonicity at $\alpha = 0$.* Since $x \mapsto x/(1 - x)$ is strictly increasing on $x < 1$, it suffices that $N$ is strictly increasing, i.e. that $\frac{d}{d\eta}\left[\eta^2\Phi'(\eta)\right] = \eta\left[2\Phi'(\eta) + \eta\Phi''(\eta)\right] > 0$. Direct computation gives

$$\begin{aligned} 2\Phi'(\eta) + \eta\Phi''(\eta) \quad &= \quad \frac{4}{(\eta^2 - 2\eta - 4)^4}\Big[\beta^2 q(\eta) + 8\beta\eta(\eta^4 + 10\eta^3 + 44\eta^2 + 48\eta + 72) \\ &\qquad + 4(3\eta^5 + 36\eta^4 + 120\eta^3 + 224\eta^2 + 128\eta + 64)\Big], \end{aligned}$$

with $q(\eta) = \eta^5 + 16\eta^4 + 32\eta^3 + 128\eta^2 - 16\eta + 64$. The denominator is a fourth power, hence positive. Every coefficient in the bracket is positive except the single term $-16\beta^2\eta$ inside $\beta^2 q(\eta)$, which is dominated because $128\eta^2 - 16\eta + 64$ has negative discriminant $(256 - 4 \cdot 128 \cdot 64 < 0)$, so $q(\eta) > 0$ for all $\eta > 0$. Hence $N$ is strictly increasing on all of $\eta > 0$, and $s/\epsilon$ is strictly increasing on the entire admissible $\gamma$-interval at $\alpha = 0$.

*Step 4: extension to $\alpha$ small.* With quadratic costs, $\gamma\Psi(\gamma; \alpha)$ is an explicit rational function of $(\gamma, \alpha)$, jointly continuous (together with its $\gamma$-derivative) wherever the binding Jacobian minor is nonzero, and the upper endpoint $\bar{\gamma}(\alpha)$ of the admissible $\gamma$-interval (the first zero of that minor) is continuous in $\alpha$. We cover $(0, \bar{\gamma}(\alpha))$ by three zones and take $\breve{\alpha}$ as the minimum of the three resulting thresholds. (i) *Near $\gamma = 0$ :* at $\alpha = 0$, $\gamma\Psi(\gamma)/C''(0) = \frac{1}{2}(\beta^2 + 4)\gamma^2 + O(\gamma^3)$ (using $\Phi'(0) = (\beta^2 + 4)/2$), so its $\gamma$-derivative is positive on some $(0, \delta]$; the quadratic coefficient is continuous in $\alpha$ (it equals $\kappa_u^0 f_x^0 + \kappa_\mu^0 f_y^0$ evaluated at the origin, which is positive at $\alpha = 0$ and continuous), and the $O(\gamma^3)$ remainder is uniform for $\alpha$ in a compact neighborhood of 0, so positivity of the derivative on $(0, \delta]$ persists for all $\alpha$ small. (ii) *Compact middle:* on $[\delta, \bar{\gamma}(\alpha) - \delta]$, Step 3 gives $d(\gamma\Psi)/d\gamma \geq m(\delta) > 0$ at $\alpha = 0$, and joint continuity extends strict positivity to all $\alpha$ below some threshold. (iii) *Near the boundary:* write $\gamma\Psi/C''(0) = P/T$ with $T$ the normalized binding minor. At $\alpha = 0$,

$T = 1 - N$ with $T' = -N' < 0$ strictly (Step 3) and $P(\bar{\gamma}(0)) = N(\bar{\gamma}(0)) = 1 > 0$, so $\frac{d}{d\gamma}(P/T) = (P'T - PT')/T^2 > 0$ on a neighborhood $(\bar{\gamma} - \delta, \bar{\gamma})$; the strict inequalities $P > 0$ and $T' < 0$ at $(\bar{\gamma}(0), 0)$ persist, with a uniform neighborhood width, for all $\alpha$ small by continuity. Combining the three zones proves the claim for all $\alpha \in [0, \breve{\alpha})$. $\square$

**Remark.** (1) The threshold $\breve{\alpha}$ is not explicit, and need not cover the entire range $\alpha < \hat{\alpha}(\epsilon)$ of part (iii); when the two ranges differ, Proposition 1(iv)'s salience clause is established only on their intersection. Some restriction on $\alpha$ is unavoidable: numerically, monotonicity fails in the admissible region for $\alpha \gtrsim 0.62$ (attainable, since $\alpha < \bar{\alpha}$ requires only $\beta \gtrsim 10\varpi$ there); for example, at $\alpha = 0.83$, $\beta = 33.7\varpi$ the full non-linear equilibrium exhibits $s^A$ single-peaked in $\gamma$. An explicit uniform bound ($\alpha \leq 1/2$) can be established by a computer-assisted positivity certificate, available on request. (2) The conditions $\sigma_1^a, \sigma_1^b \in (0, 1)$ are not implied by quadratic costs, which impose no barrier as $s^A$ approaches $\pm(1/2 - \epsilon)$. For every $\alpha$ and $\beta > 0$ they hold throughout the admissible $\gamma$-range once $\epsilon$ is below a strictly positive threshold $\bar{\epsilon}(\alpha, \beta)$; this threshold is largest precisely in the small-$\alpha$ range covered by the lemma (numerically $\bar{\epsilon} \approx 0.04$ at $\alpha = 0.1$, $\beta = 2\varpi$), so the interiority qualifier is least restrictive exactly where the lemma applies.

# B Choice of fixed covariate penalty $\psi$

Following GST, the fixed covariate cost should be chosen so that, when the populist penalty is set to an almost negligible level for all bigrams ($\lambda_j = 10^{-10}$), the resulting partisanship estimates are similar to those obtained under MLE. As discussed in the main text, the role of $\lambda_j$ in the main specification is to reduce the finite-sample bias of the estimator. Therefore, we want the fixed covariate cost $\psi$ calibrated such that, when we do not attempt to reduce finite-sample bias, the estimates resemble the MLE results. At the same time, we prefer to keep the fixed covariate cost relatively high, so that the covariates do not absorb too much of the variation in bigram usage.

To find an optimal value, we explore different fixed cost values – $\psi \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$. Both $10^{-5}$ and $10^{-6}$ lead to a pattern of partisanship similar to the MLE (results available upon request). We thus choose $10^{-5}$ as a fixed cost for the rest of the paper.

# C Partisanship with four political types

Suppose there are four political types, defined on two dimensions: populist / non-populist and left / right. Hence, rewrite (7) as:

$$U_{ijt}^{P_iR_i} = \alpha_{jt} + X_{it}\gamma_j + \phi_{jt}P_i + \psi_{jt}R_i$$

with $P_i = 0, 1$, $R_i = 0, 1$. We thus have four political types, Populist - Right, Non-Populist - Right, and so on, and estimate a vector of probabilities for each type, $q_{\mathbf{it}}^{\mathbf{P}_i\mathbf{R}_i} = \left\{q_{ijt}^{P_iR_i}\right\}$.

The posterior probability that $i$ is of type, say, Populist - Right, with neutral priors and conditional on bigram $j$, is:

$$\hat{\rho}_{ijt}^{11} = \frac{\hat{q}_{ijt}^{11}}{\hat{q}_{ijt}^{11} + \hat{q}_{ijt}^{10} + \hat{q}_{ijt}^{01} + \hat{q}_{ijt}^{00}}$$

and similarly for the other types.

Partisanship of $i$ in this more granular model is the weighted average of these posterior probabilities across all bigrams for all possible types, weighted by the probability of their occurrence, namely:

$$\hat{\pi}_{it4} = \frac{1}{4}\sum\nolimits_j \sum_{P=0,1;R=0,1} \hat{q}_{ijt}^{PR}\hat{\rho}_{ijt}^{PR}$$

This measures the predictability of $i's$ (granular) type by averaging his true type $P_iR_i$ with all his possible "clones" with neutral priors. If all types always use the same bigrams, then they are indistinguishable and $\hat{\pi}_{it4} = 1/4$. If instead they always use different bigrams, then $\hat{\pi}_{it4} = 1$.

We can also estimate the predictability of $P$ vs $NP$, when there are 4 political types as:

$$\hat{\pi}_{it4}^{P/NP} = \frac{1}{4}\sum\nolimits_j [(\hat{q}_{ijt}^{11} + \hat{q}_{ijt}^{10})\hat{\rho}_{ijt}^{1\circ} + (\hat{q}_{ijt}^{01} + \hat{q}_{ijt}^{00})(1 - \hat{\rho}_{ijt}^{1\circ}]$$

where $\hat{\rho}_{ijt}^{1\circ} = \hat{\rho}_{ijt}^{11} + \hat{\rho}_{ijt}^{10}$ is the posterior that $i$ is populist ($P_i = 1$). This expression for partisanship measures the average predictability of $i's$ type being $P$ or $NP$, when there are four possible political types. Similarly, we can define right-left partisanship when there are four political types as:

$$\hat{\pi}_{it4}^{R/L} = \frac{1}{4}\sum\nolimits_j [(\hat{q}_{ijt}^{11} + \hat{q}_{ijt}^{01})\hat{\rho}_{ijt}^{\circ 1} + (\hat{q}_{ijt}^{10} + \hat{q}_{ijt}^{00})(1 - \hat{\rho}_{ijt}^{\circ 1}]$$

where now $\hat{\rho}_{ijt}^{\circ 1} = \hat{\rho}_{ijt}^{11} + \hat{\rho}_{ijt}^{01}$.

By Bayes rule, the posterior that the speaker has features $X_{it}$, rather than being some other politician of the same country $c$, conditional on observing bigram $j$ in quarter $t$, starting with neutral prior $(1/n_c)$, is:

$$\hat{\rho}^c_{ijt} = \frac{\hat{q}_{ijt}}{\sum_{i\in c}\hat{q}_{ijt}} = \frac{\hat{q}_{ijt}}{n_c\hat{q}_{cjt}}$$

Define partisanship $\hat{\pi}^c_{it}$ as the weighted average of $\hat{\rho}^c_{ijt}$ across bigrams

$$\begin{aligned}
\hat{\pi}^c_{it} &= \sum_j \hat{q}_{ijt}\hat{\rho}^c_{ijt} \\
&= \frac{1}{n_c}\sum_j \frac{\hat{q}^2_{ij}}{\hat{q}_{cjt}} \\
&= \frac{1}{n_c}[1+\hat{\chi}^2_{ict}]
\end{aligned}$$

where for the last step we used:

$$\begin{aligned}
\hat{\chi}^2_{ict} &= \sum_j \frac{\hat{q}^2_{ijt}+\hat{q}_{cjt}{}^2-2\hat{q}_{ijt}\hat{q}_{cjt}}{\hat{q}_{cjt}} \\
&= \sum_j \frac{\hat{q}^2_{ijt}}{\hat{q}_{cjt}} + \sum_j \hat{q}_{cjt} - 2\sum_j \hat{q}_{ijt} \\
&= \sum_j \frac{\hat{q}^2_{ijt}}{\hat{q}_{cjt}} - 1
\end{aligned}$$

# D Construction of confidence intervals

Given that standard nonparametric bootstrap is known to be invalid for lasso regression (Chatterjee and Lahiri, 2011), confidence intervals are estimated via subsampling, similar to the way it is done in GST. To do so, we randomly draw speakers without replacement to create 100 subsamples each containing (up to integer restrictions) one-fifth of all speakers and, for each subsample $k$, compute the penalized estimate of partisanship $\pi^k_t$. The confidence interval (90%) is then constructed according to the following formula around the estimated $\pi^*_t$:

$$CI(\pi^*_t) = \frac{1}{2} + [\exp(\log(\pi^*_t - 1/2) - Q^{*k}_{t(95)}/\sqrt{\tau}); \exp(\log(\pi^*_t - 1/2) - Q^{*k}_{t(5)}/\sqrt{\tau})],$$

where $\tau$ is the number of speakers in the full sample, $Q^{*k}_{t(b)}$ is the $b$th order statistic of:

$$Q^{*k}_t = \sqrt{\tau_k}(\log(\pi^k_t - 1/2) - \log([\frac{1}{100}\sum_{l=1}^{100}\pi^l_t] - 1/2))$$

# E Speech informativeness

Following the logic of the model in Section 3, and as in GST, for each politician $i$ and calendar quarter $t$ we randomly draw $N$ bigrams from a multinomial distribution with bigram choice probabilities given by the estimated frequencies, $\hat{\mathbf{q}}_{it}^{P_i} = \{\hat{q}_{ijt}^{P_i}\}$ ($P_i$ denotes politician $i$'s true type). Let $\hat{q}_{int}^{P_i}$ denote the frequency of use of the $n$-th drawn bigram for politician $i$ in calendar quarter $t$. Then, for any sequence length $n \in \{1, 2, \ldots, N\}$, we compute the following measure of speech informativeness conditional on the observed sequence:

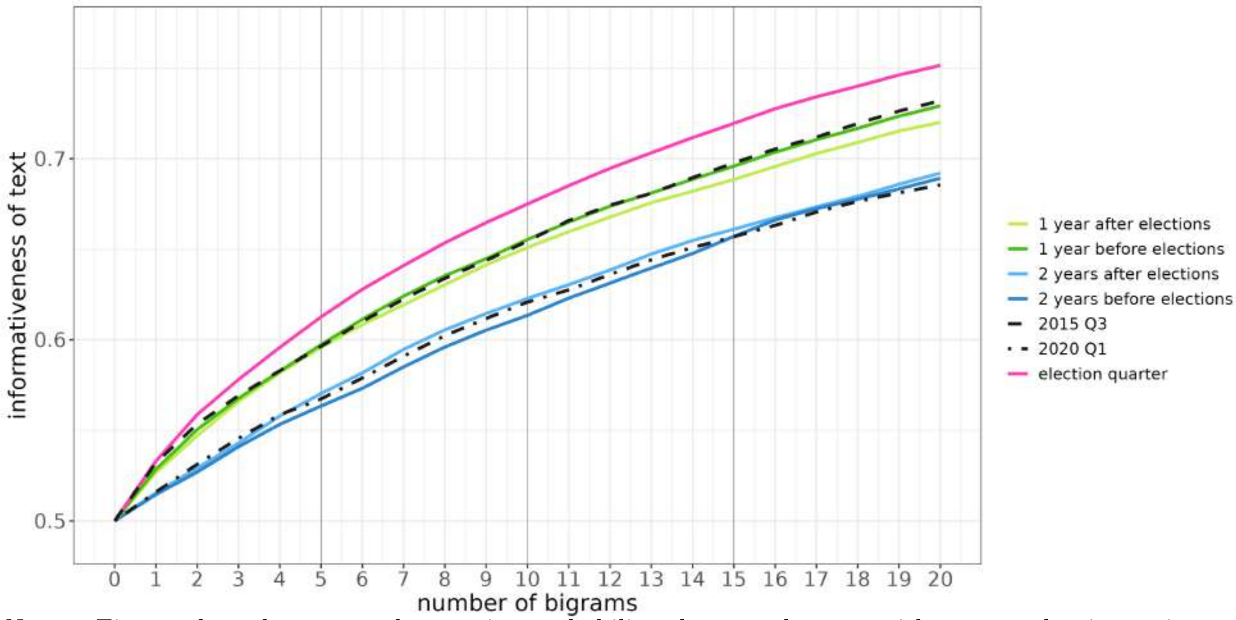
$$\eta_{itn} = \frac{\hat{q}_{int}^{P_i} \eta_{itn-1}}{\hat{q}_{int}^{P_i} \eta_{itn-1} + \hat{q}_{int}^{1-P_i}(1 - \eta_{itn-1})}, \tag{E1}$$

where $\eta_{it0} = \frac{1}{2}$ for all $i$ and calendar quarters $t$ (i.e. as before, an observer starts with a neutral prior).[1] Averaging $\eta_{itn}$ over politicians $i$ in a given quarter $t$ we obtain a measure of speech informativeness $\eta_{tn}$. To obtain a similar measure at a given quarter distance $d$ from the election, we take all the politicians-quarter observations at a given quarter distance $d$, and average $\eta_{id(i,t)n}$ over $i$ to obtain $\eta_{dn}$.

Figure E1 presents the average results of repeating this procedure 10 times through a Monte Carlo simulation *for each politician* in each quarter. The vertical axis measures speech informativeness about the politician being populist or non-populist, the horizontal axis measures the number of bigrams. One tweet corresponds to about 5 bigrams (the average length of a bigram is about 20 characters and the average length of a tweet is about 100 characters, although tweets also include many rarely used bigrams that we have excluded from the analysis). Different colors correspond to different distances from the election date. The figure yields the estimates described in the text.

[1] This measure corresponds to the expected posterior probability of correctly identifying politician $i$'s type after observing $n$ bigrams, starting with a neutral prior.

Figure E1: Expected posterior belief of an observer with a neutral prior after reading a given number of bigrams




*Notes.* Figure plots the expected posterior probability that an observer with a neutral prior assigns to politician $i$'s true populist affiliation after hearing $n$ bigrams ($\eta_{itn}$), averaged over $i$. $\eta_{itn}$ is defined as in Equation E1. Estimated bigram-usage probabilities $\hat{\mathbf{q}}_{it}$ from our benchmark specification are used. Vertical lines indicate approximately one, two, and three tweets (five, ten, and fifteen bigrams, respectively).

# F Data

## F.1 Classification of Populist Leaders

### F.1.1 Switching political types

We directly asked ChatGPT whether the politician changed between being a populist and a non-populist using the prompt below, and repeated the process 4 times. Over the 367 politicians in our sample, the answer is "Unchanged" for 363 of them with full consistency (4/4). We just have 3 cases where ChatGPT answered 2 times over 4 "Changed from NP to P" or "from P to NP". These politicians are Cabo Daciolo from Brazil, Mark Latham from Australia (both from NP to P), and Juliette Boulet from Belgium (from P to NP). For Paweł Piotr Kukiz 3 out of 4 times ChatGPT gave an answer "Changed from NP to P".

The ChatGPT prompt used to assess whether politicians switched from one political type to another was as follows:

*I will give you the name of a present or past leader of a political party and the country they are from. I am interested in whether this person **changed** between being a populist* *and a non-populist, or vice-versa, during the period 2010–2022.*

*My definition of populism is the following: A leader is defined as populist if he or she divides society into two artificial groups – "the people" vs. "the elites" – and then claims to be the sole representative of the true people. Populists place the alleged struggle of the people ("us") against the elites ("them") at the center of their political campaign and governing style. More precisely, populists typically depict "the people" as a suffering, inherently good, virtuous, authentic, ordinary, and common majority, whose collective will is incarnated in the populist leader. By contrast, "the elite" is an inherently corrupt, self-serving, power-hoarding minority, negatively defined as all those who are not "the people".*

*Determine whether the leader **changed** between being a populist and a non-populist, or vice-versa, **between 2010 and 2022** and give me one of these four possible answers:*

*1. **Changed from P to NP**: if the answer is yes and the politician changed from being a Populist to a Non-Populist.*

*2. **Changed from NP to P**: if the answer is yes and the politician changed from being a Non-Populist to a Populist.*

*3. **Unchanged**: if the answer is no, meaning the politician was either always a Populist or always a Non-Populist.*

*4. **Ambiguous**: if the answer is ambiguous because either you have lack of sufficient information or it is truly an ambiguous case.*

*Rules: - Output only a single label. - Be concise and accurate.*

### F.1.2 Comparison of our classification with fully automated ChatGPT classification

We ran an exercise similar to the one above to validate our manual classification of populist politicians. To do so, we asked ChatGPT to classify politicians as either populists, non-populists or ambiguous. The results show a total of 26 cases of disagreement between the fully automated ChatGPT classification and the benchmark classification used in the main text (the classification in which the remaining ambiguous cases are considered populists). In all 26 cases, our benchmark classification identifies the politicians as populists, whereas the fully automated ChatGPT classification labels them as non-populists. Of these 26 politicians, 10 are classified as populists in our data because they belong to a populist party, following our approach for ambiguous cases to reduce the overall number of ambiguities.

Among the remaining 16 politicians, 7 are classified as ambiguous and do not belong to a populist party; therefore, following the approach described in the main text, we treat them as populists in the benchmark specification. The other 9 politicians are considered populists in our classification but are labeled as non-populists by the automated ChatGPT classification. These politicians are Beata Maria Kuśińska, Catarina Martins, Elizabeth Warren, Jeremy Corbyn, Judit Varga, Kristian Thulesen Dahl, Pia Kjærsgaard, Roberto Maroni, and Zoltán Kovács.

## F.2 Fixation Index

To validate the classification of $P$ vs $NP$ leaders we measured its performance vis-a-vis alternative classifications. For each politician, Claude 4.6 Sonnet was asked to report a Populist Score by answering the following question: *"I have a sample of politicians and I want you to answer how certain you are that the politician is populist. 0% means definitely non-populist, 100% means definitely populist."* Once this task was completed, partitions of the sample into two groups were created. First, we reconstructed the benchmark classification used in the paper, which has 84 Populists and 283 Non-Populists. A second partition was created by taking the 84 politicians with the highest Populist Score as estimated by Claude as one group. We then computed 150 more random partitions, maintaining the group sizes of 84 and 283.

For each partition $m$, a $\chi^2$-based measure of the Fixation Index was computed at the quarter level and then averaged across all quarters. Probabilities were estimated with the standard specification, including all covariates. The group dummy was not penalized but included among the covariates and interacted with calendar quarter.

Consider the usual notation, where $q_{i,j,t}$ denotes the usage probability of bigram $j$ by individual $i$ in calendar quarter $t$. A partition $m$ splits the sample into two groups $g \in \{P, NP\}$, and we let $g_m(i)$ denote the group to which individual $i$ is assigned under partition $m$. Define the overall mean for each bigram,

$$q_{j,t} = \frac{1}{N_t} \sum_i q_{i,j,t},$$

and the group-specific means for each bigram within each quarter,

$$q_{g,j,t} = \frac{1}{N_{g,t}} \sum_{i \,:\, g_m(i)=g} q_{i,j,t},$$

where $N_t$ is the number of individuals in quarter $t$ and $N_{g,t}$ is the number of individuals in group $g$ in quarter $t$. We measure total and between-group heterogeneity in quarter $t$ respectively by:

$$\chi^2_t = \frac{1}{N_t} \sum_i \sum_j \frac{(q_{i,j,t} - q_{j,t})^2}{q_{j,t}}.$$

$$\chi^2_{m,t} = \frac{1}{N_t} \sum_{g \in \{P,NP\}} N_{g,t} \sum_j \frac{(q_{g,j,t} - q_{j,t})^2}{q_{j,t}},$$

The Fixation Index was then computed as

$$F_{m,t} = \frac{\chi^2_{m,t}}{\chi^2_t},$$

and the average for partition $m$ is

$$\hat{F}_m = \frac{1}{T} \sum_{t=1}^{T} F_{m,t}.$$

This measure has a clear interpretation: the share of variation in predicted bigram usage due to between-group variation, for the groups defined by partition $m$.

Figure F1 plots the distribution of $\hat{F}_m$ over the generated partitions of politicians. The left-most vertical line corresponds to the Claude partition, the right-most vertical line to our benchmark classification. Thus, our classification uncovers the latent language patterns in the texts very well, and it outperforms the partition based on the Claude classification.

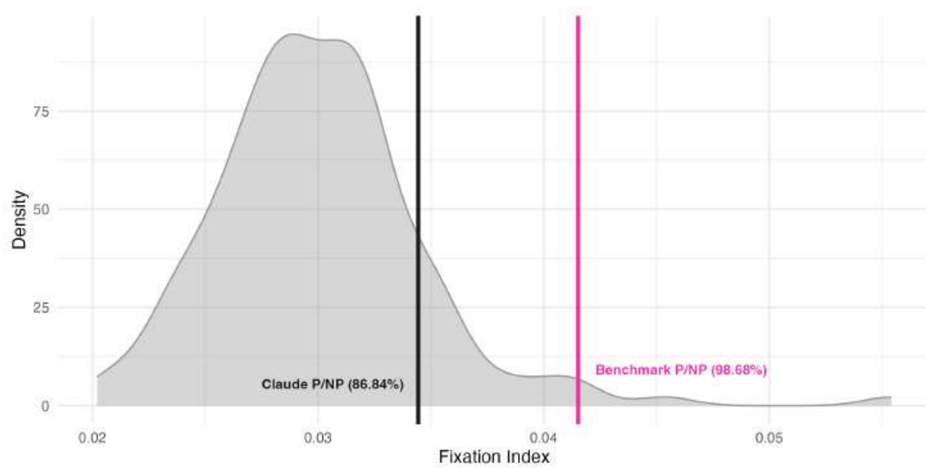



Figure F1: Fixation index distribution over generated partitions of politicians.

*Notes. Benchmark P/NP* refers to the classification used as the benchmark in the paper. *Claude P/NP* refers to a classification that partitions the sample into populist and non-populist groups of the same size as in the benchmark, based on populist scores generated by Claude Sonnet.

## F.3 Classification of Niche Parties

To classify niche parties, the following prompt was used:

You are a helpful research assistant. I will give you the name of a past or present political party and what country they are from. I am interested in whether they are a 'niche' political party. In particular, a party is niche if it satisfies all three of the following criteria:

1) Niche parties reject the traditional class-based orientation of politics. So instead of prioritizing demands for more or less redistribution, these parties politicize other issues which were previously not represented in party competition. In many cases, green parties or radical right parties would be examples of this.

2) The issues raised by niche parties are not only novel, but they often do not coincide with existing lines of political division. Niche parties appeal to groups of voters that may cross-cut traditional partisan alignments. As a result, cases of voter defection between "unlikely" party pairs have occurred. The defection of former British Conservative voters to the Green Party in 1989 and former French Communist party voters to the radical right Front National in 1986 are typical examples.

3) Niche parties keep a very narrow policy agenda, focusing only on a few key policies. So for example the Green Party in Germany is not a niche party because it is a mainstream party with a very broad policy agenda.

Respond with 'yes' if the party is niche and 'no' if it is not.

## F.4 International connections of Leaders

We explore cross-border connections of political leaders along the P/NP and L/R divides using data on direct followers for 305 (out of 367) leaders across the 21 countries covered by our study. Our goal is to evaluate differential rates of connection between groups, and to do so we adopt a dyadic regression approach similar to Colonnelli, Neto and Teso (2025). This approach allows us to account for group imbalances, as well as to control for confounding factors.

We transform our directed network into a dataset of cross-border dyads: ordered pairs $(i, j)$ in which politician $i$ (the sender) may follow politician $j$ (the receiver) located in a different country. Each dyad is thus a potential tie, realised when $i$ actually follows $j$; the unit of observation is the dyad. The full dataset of cross-border dyads comprises 88,102 potential ties, of which 2,523 are realized ties. We define homophily as the tendency for cross-border ties to occur disproportionately between politicians of the same type. Using a linear probability model, we compare the probability that a same-type dyad forms a tie with the corresponding probability for a different-type dyad. We then test whether this same-type selectivity is stronger among populist senders than among non-populist senders, and we repeat the same analysis for the Left/Right divide.

We estimate the following linear probability model over all ordered cross-border dyads:

$$\mathrm{Tie}_{ij} = \alpha + \beta_1\,\mathrm{SameType}_{ij} + \beta_2\,\mathrm{Pop}_i + \beta_3\left(\mathrm{SameType}_{ij} \times \mathrm{Pop}_i\right) + \mathbf{X}'_{ij}\gamma + \mu_{c(i)} + \nu_{c(j)} + \varepsilon_{ij}, \quad \text{(F1)}$$

where $i$ is the potential sender (follower), $j$ the potential receiver (followed). The dependent variable $\mathrm{Tie}_{ij} = \mathbf{1}\{i \text{ follows } j\}$ equals one when the tie is realized and zero otherwise. The regressor $\mathrm{SameType}_{ij} = \mathbf{1}\{\mathrm{Pop}_i = \mathrm{Pop}_j\}$ indicates that sender and receiver share the same populism status, and $\mathrm{Pop}_i = \mathbf{1}\{i \text{ is populist}\}$ indicates a populist sender. The vector $\mathbf{X}'_{ij}$ contains dyad-level indicators for shared gender and shared left–right classification, together with sender and receiver gender, birth year, right-wing status, and indicators for having held a position as head of government or head of state (as proxy for leader popularity). The $\mu_{c(i)}$ and $\nu_{c(j)}$ terms are sender- and receiver-country fixed effects. They absorb differences across sender and receiver countries in the probability of forming a cross-border tie. In the sender-fixed-effects specification, $\mu_{c(i)}$ and all sender-level covariates are absorbed by an individual sender effect $\phi_i$.[2] Throughout, standard errors are clustered two-way on sender and receiver identity to account for the dependence among all dyads that share a sender and all dyads that share a receiver. We repeat the same analysis for the Left-Right divide.

The coefficient $\beta_1$ measures the increase in tie probability associated to same type $NP \rightarrow NP$ match, compared to a $NP \rightarrow P$ match (non-populist homophily); $\beta_2$ is the difference in tie probability for a $P \rightarrow NP$ match compared to a $NP \rightarrow P$ match. Finally, $\beta_3$ measures the differential homophily of populist compared to non-populist leaders, which estimates the difference in tie probability associated to a $P \rightarrow P$ match (wrt a $P \rightarrow NP$ match), compared to a $NP \rightarrow NP$ match (wrt an $NP \rightarrow P$ match).

Table F1 shows the results of our cross-border estimates. In the unconditional specification, the coefficient $\beta_1$ is positive and significant, while $\beta_3$ is positive, but not statistically significant. $\beta_3$ becomes larger and statistically significant controlling for sender-country and receiver-country fixed effects, as well as dummies for Head of Government or Head of State positions. In these specifications populists appear to be more homophilic in their cross-border connections compared to non-populists, and the increase in tie probability associated with a same-type match is around 2.3 percentage points larger among populist senders than among non-populist senders.

We repeat the same analysis along the Left/Right divide in Table F2. We do not find a differential right-wing (vs left-wing) homophily in cross-border ties. The coefficient $\beta_3$ is statistically significant only when we exclude from the sample politicians leaning to a central position in the political spectrum (having a RILE index in the (-5,5) interval, Column 4).

[2] We do not include receiver fixed effects, as they would absorb the receiver's populism status, leaving $\beta_1$ and $\beta_3$ no longer separately identified.

Table F1: Differential populist homophily in cross-border political ties

| | (1) | (2) | (3) |
|---|---|---|---|
| $(NP \to NP) - (NP \to P)$ $[\beta_1]$ | 0.015** | 0.006 | 0.006 |
| | (0.007) | (0.005) | (0.005) |
| Differential homophily $[\beta_3]$ | 0.006 | 0.023** | 0.023** |
| | (0.012) | (0.010) | (0.010) |
| Sender-country FE | No | Yes | Absorbed |
| Receiver-country FE | No | Yes | Yes |
| Sender individual FE | No | No | Yes |
| Individual-level controls | No | Yes | Yes |
| HoG/S controls | No | Yes | Yes |
| Observations | 88,102 | 88,102 | 88,102 |
| Adjusted $R^2$ | 0.003 | 0.061 | 0.097 |
| Sender/receiver clusters | 305/305 | 305/305 | 305/305 |

*Notes:* The sample contains all ordered cross-border dyads. The dependent variable equals one if sender $i$ follows receiver $j$. Differential homophily is $\beta_3 = \left[\Pr(P \to P) - \Pr(P \to NP)\right] - \left[\Pr(NP \to NP) - \Pr(NP \to P)\right]$. Individual-level controls include shared gender, shared right-wing status, and sender and receiver gender, birth year, and right-wing status. HoG/S controls indicate whether the sender and receiver have held a position as head of government or head of state. Sender-level covariates are absorbed by sender fixed effects in column (3). Standard errors, two-way clustered by sender and receiver, are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table F2: Differential right-wing homophily in cross-border political ties

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $(L \to L) - (L \to R)$ $[\beta_1]$ | 0.011* | 0.017** | 0.017** | 0.026*** |
| | (0.006) | (0.007) | (0.007) | (0.009) |
| Differential homophily $[\beta_3]$ | $-0.002$ | $-0.016$ | $-0.017$ | $-0.031$* |
| | (0.012) | (0.013) | (0.013) | (0.017) |
| Sender-country FE | No | Yes | Absorbed | Absorbed |
| Receiver-country FE | No | Yes | Yes | Yes |
| Sender individual FE | No | No | Yes | Yes |
| Individual-level controls | No | Yes | Yes | Yes |
| HoG/S controls | No | Yes | Yes | Yes |
| Drops $\lvert\text{RILE}\rvert < 5$ | No | No | No | Yes |
| Observations | 88,102 | 88,102 | 88,102 | 59,222 |
| Adjusted $R^2$ | 0.001 | 0.061 | 0.097 | 0.091 |
| Sender/receiver clusters | 305/305 | 305/305 | 305/305 | 250/250 |

*Notes:* The sample contains all ordered cross-border dyads. The dependent variable equals one if sender $i$ follows receiver $j$. A politician is classified as right-wing ($R$) if the RILE (right–left) score of their party's manifesto is positive, and non-right ($L$) otherwise. Differential homophily is $\beta_3 = \left[\Pr(R \to R) - \Pr(R \to L)\right] - \left[\Pr(L \to L) - \Pr(L \to R)\right]$. Individual-level controls include shared gender, shared populism status, and sender and receiver gender, birth year, and populism status. HoG/S controls indicate whether the sender and receiver have held a position as head of government or head of state. Sender-level covariates are absorbed by sender fixed effects in columns (3) and (4). Column (4) repeats the specification of column (3) on the subsample that excludes politicians whose RILE score lies in $(-5, 5)$. Standard errors, two-way clustered by sender and receiver, are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Differences in homophily over international connections appear to be stronger along the P/NP divide than along the L/R divide when we compare the two differential effects within the same regression (see Table F3). In particular we run the following regression:

$$\text{Tie}_{ij} = \alpha + \delta_1\,\text{SameType}_{ij} + \delta_2\,\text{SameRight}_{ij} + \mathbf{X}'_{ij}\boldsymbol{\gamma} + \mu_{c(i)} + \nu_{c(j)} + \varepsilon_{ij}, \tag{F2}$$

where $\text{SameType}_{ij} = \mathbf{1}\{\text{Pop}_i = \text{Pop}_j\}$ and $\text{SameRight}_{ij} = \mathbf{1}\{\text{Right}_i = \text{Right}_j\}$ indicates that the sender and receiver have the same political ideology (P or NP and L or R respectively). The vector $\mathbf{X}'_{ij}$ now contains main effects for receiver and sender right-wing and populist status, along with dyad-level indicators for shared gender, and receiver and sender gender and year of birth. All other terms are as in Equation F1.

The results are reported in Table F3. The benchmark specification (4) shows that sharing populist status increases the probability of a cross-border tie by 1.8 percentage points, while sharing left/right-wing alignment increases it by 0.9 percentage points. The difference between these two probabilities is around 1 percentage point and statistically significant at the 5% level. This difference becomes smaller and not statistically significant when politicians close to the ideological center are excluded from the sample, specification (5).

Table F3: P/NP versus L/R homophily in cross-border political ties

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| P/NP homophily [$\delta_1$] | 0.017*** | 0.017*** | 0.018*** | 0.018*** | 0.016*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| L/R homophily [$\delta_2$] | 0.009*** | 0.009*** | 0.009*** | 0.009*** | 0.011*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Difference ($\delta_1 - \delta_2$) | 0.009** | 0.009** | 0.009** | 0.009** | 0.006 |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Sender-country FE | No | Yes | Yes | Absorbed | Absorbed |
| Receiver-country FE | No | Yes | Yes | Yes | Yes |
| Sender individual FE | No | No | No | Yes | Yes |
| Demographic controls | No | Yes | Yes | Yes | Yes |
| HoG/S controls | No | No | Yes | Yes | Yes |
| Drops $\lvert\text{RILE}\rvert < 5$ | No | No | No | No | Yes |
| Observations | 88,102 | 88,102 | 88,102 | 88,102 | 59,222 |
| Adjusted $R^2$ | 0.004 | 0.040 | 0.061 | 0.097 | 0.091 |
| Sender/receiver clusters | 305/305 | 305/305 | 305/305 | 305/305 | 250/250 |

*Notes:* The sample contains all ordered cross-border dyads. The dependent variable equals one if sender $i$ follows receiver $j$. Row [$\delta_1$] reports the coefficient on an indicator for sender and receiver sharing the same populism status; row [$\delta_2$] the coefficient on an indicator for sharing the same left–right status, where a politician is classified as right-wing if the RILE score of their party's manifesto is positive. Both indicators enter the same regression, each alongside its own sender and receiver main effects (included in all columns). The reported difference tests whether populist homophily exceeds left–right homophily. Demographic controls include shared gender and sender and receiver gender and birth year. HoG/S controls indicate whether the sender and receiver have held a position as head of government or head of state. Sender-level covariates are absorbed by sender fixed effects in columns (4) and (5). Column (5) excludes politicians whose RILE score lies in $(-5, 5)$. Standard errors, two-way clustered by sender and receiver, are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

## F.5 Classification of Leaders along Left-Right

### F.5.1 ChatGPT Classification

Classification of politicians with ChatGPT was performed using the model GPT-4o-mini and with the following prompt:

*I will give you the name of a present or past leader of a political party and what country they are from. I am interested in whether they are a left-wing or right-wing politician in the period 2010-2021. In particular, give me one of these three possible answers:*

*1. left: commonly considered a left-wing leader 2. right: commonly considered a right-wing leader 3. ambiguous: ambiguous whether they are left- or right-wing because either: (i) they are truly an ambiguous case (ii) they change from being left- to right-wing (or vice-versa) during the period (iii) you do not have enough information about the individual to give an answer*

*Rules: - Output only a single label. - If the label is ambiguous, specify whether it is because of (i), (ii) or (iii). - Be concise and accurate.*

ChatGPT classification exhibits some randomness, therefore results slightly change when the same procedure is repeated. Reassuringly, in the 4 iterations there was no case in which ChatGPT switched a politician's classification from right to left or vice versa; the classification only switched between right and ambiguous, or between left and ambiguous.

To account for this, the classification has been repeated 4 times, and politicians have been classified as "Right" if ChatGPT classified the politician in this category all 4 times, the same with "Left". Otherwise, politicians have been assigned to the "ambiguous" category. This process led to 82 ambiguous cases (later 64 of these politicians were classified as left by the RILE classification, 18 as right). The number of politicians which were positively classified is 285: the confusion matrix for this sample against the RILE classification is reported below – Table F4.

### F.5.2 RILE Classification

We also classify all the politicians using the party-level classification. The Manifesto Project Dataset measures left-right positions of political parties through the 'rile' index, which measures how much a party talks about left-wing or right-wing issues. We consider, for each party, the most recent codification of this index and classify politicians as right if they belong to a party whose RILE score is greater than zero, as left otherwise. Politicians that have no match with the dataset have been manually classified (with the use of ChatGPT).

We prefer the ChatGPT classification to the one based on the Manifesto Project Dataset, as we believe that the former could better capture within-party factions which are relevant in classifying individual leaders, and it is more comparable internationally. In Appendix Figure H18 we present our main results for the L-R classification based on the RILE index instead of ChatGPT.

Table F4: Left-Right Classifications Confusion Matrix - Absolute Values

| | GPT Left | GPT Right | |
|---|---|---|---|
| RILE Left | 147 | 46 | 193 |
| RILE Right | 12 | 80 | 92 |
| | 159 | 126 | 285 |

## F.6 Summary of leaders classifications

**Australia**

Andrew Bartlett (NP, L), Bill Shorten (NP, L), Bob Katter (P, R), Campbell Newman (NP, R), Christine Milne (NP, L), Clive Palmer (P, R), Deb Frecklington (NP, R), John Anderson (NP, L), John-Paul Langbroek (NP, R), Julia Gillard (NP, L), Kevin Rudd (NP, L), Kim Beazley (NP, L), Malcolm Turnbull (NP, R), Mark Latham (P, R), Natasha Stott Despoja (NP, L), Nick Xenophon (P, L), Pauline Hanson (P, R), Richard Di Natale (NP, L), Scott Morrison (NP, R), Tim Nicholls (NP, R), Tony Abbott (P, R)

**Austria**

Alexander Van der Bellen (NP, L), Beate Meinl-Reisinger (NP, L), Christian Kern (NP, L), Heinz-Christian Strache (P, R), Josef Bucher (P, R), Matthias Strolz (NP, L), Norbert Hofer (P, R), Pamela Rendi-Wagner (NP, L), Peter Pilz (NP, L), Sebastian Kurz (P, R), Ulrike Lunacek (NP, L), Werner Kogler (NP, L)

**Belgium**

Alexander De Croo (NP, R), Bart De Wever (P, R), Benoît Lutgen (NP, R), Caroline Gennez (NP, L), Charles Michel (NP, R), Didier Reynders (NP, L), Elio Di Rupo (NP, L), Guy Verhofstadt (NP, L), Gwendolyn Rutten (NP, R), Herman Van Rompuy (NP, L), Jean-Marc Nollet (NP, L), John Crombez (NP, L), Joëlle Milquet (NP, R), Juliette Boulet (NP, L), Marianne Thyssen (NP, R), Meyrem Almaci (NP, L), Peter Mertens (P, L), Philippe Henry (NP, L), Philippe Lamberts (NP, L), Stefaan De Clerck (NP, R), Tom Van Grieken (P, R), Wouter Beke (NP, R), Wouter Van Besien (NP, L), Yves Leterme (NP, L), Zoé Genot (NP, L)

**Brazil**

Anthony Garotinho (P, L), Aécio Neves (NP, R), Cabo Daciolo (P, R), Ciro Gomes (P, L), Cristovam Buarque (NP, L), Dilma Rousseff (P, L), Fernando Haddad (NP, L), Geraldo Alckmin (NP, R), Heloísa Helena (P, L), Henrique Meirelles (NP, R), Jair Bolsonaro (P, R), José Serra (NP, R), João Amoêdo (NP, R), Luciana Genro (NP, L), Luiz Inácio Lula da Silva (P, L), Marina Silva (NP, L), Rui Costa Pimenta (NP, L)

**Czech Republic**

Andrej Babiš (P, L), Ivan Bartoš (NP, L), Jan Farský (NP, L), Jiří Paroubek (NP, L), Jiří Rusnok (NP, R), Karel Schwarzenberg (NP, L), Lubomír Zaorálek (NP, L), Mirek Topolánek (NP, R), Miroslav Kalousek (NP, R), Miroslava Němcová (NP, R), Pavel Bělobrádek (NP, R), Petr Fiala (NP, R), Tomio Okamura (P, R), Vladimír Špidla (NP, L), Vojtěch Filip (NP, L)

**Denmark**

Anders Fogh Rasmussen (NP, R), Anders Samuelsen (NP, R), Bendt Bendtsen (NP, R), Helle Thorning-Schmidt (NP, L), Holger K. Nielsen (NP, L), Kristian Thulesen Dahl (P, R), Lars Barfoed (NP, R), Lars Løkke Rasmussen (NP, L), Margrethe Vestager (NP, L), Marianne Jelved (NP, L), Mogens Lykketoft (NP, L), Morten Østergaard (NP, L), Pernille Skipper (P, L), Pernille Vermund (P, R), Pia Kjærsgaard (P, R), Pia Olsen Dyhr (NP, L), Søren Pape Poulsen (NP, R), Uffe Elbæk (NP, L)

**Finland**

Alexander Stubb (NP, R), Anna-Maja Henriksson (NP, L), Anneli Jäätteenmäki (NP, L), Anni Sinnemäki (NP, L), Antti Rinne (NP, L), Carl Haglund (NP, L), Eero Heinäluoma (NP, L), Harry Harkimo (NP, L), Jussi Halla-aho (P, R), Jutta Urpilainen (NP, L), Jyrki Katainen (NP, R), Li Andersson (NP, L), Mari Kiviniemi (NP, R), Matti Vanhanen (NP, R), Osmo Soininvaara (NP, L), Paavo Arhinmäki (NP, L), Pekka Haavisto (NP, L), Petteri Orpo (NP, R), Päivi Räsänen (NP, R), Sari Essayah (NP, R), Suvi-Anne Siimes (NP, L), Tarja Cronberg (NP, L), Ville Niinistö (NP, L)

**France**

Benoît Hamon (NP, L), Bernard Cazeneuve (NP, L), Dominique Voynet (NP, L), Dominique de Villepin (NP, R), Edouard Philippe (NP, R), Emmanuel Macron (NP, L), Eva Joly (NP, L), François Bayrou (NP, L), François Fillon (NP, R), François Hollande (NP, L), Jean-Luc Mélenchon (P, L), Jean-Marc Ayrault (NP, L), Jean-Marie Le Pen (P, R), Jean-Pierre Raffarin

(NP, R), Manuel Valls (NP, L), Marine Le Pen (P, R), Nicolas Dupont-Aignan (P, R), Nicolas Sarkozy (NP, R), Noël Mamère (NP, L), Ségolène Royal (NP, L)

**Germany**

Alice Weidel (P, R), Bernd Lucke (P, R), Cem Özdemir (NP, L), Christian Lindner (NP, R), Dietmar Bartsch (NP, L), Gabi Zimmer (NP, L), Gregor Gysi (NP, L), Jürgen Trittin (NP, L), Katrin Göring-Eckardt (NP, L), Martin Schulz (NP, L), Peer Steinbrück (NP, L), Renate Künast (NP, L), Sahra Wagenknecht (P, L)

**Hungary**

Bocskor Andrea (NP, R), Deli Andor (NP, L), Deutsch Tamás (P, R), Edina Toth (P, R), Gergely Karácsony (NP, L), Gordon Bajnai (NP, L), István Ujhelyi (NP, L), Judit Varga (P, R), Katalin Novák (P, R), László Andor (NP, L), Márton Gyöngyösi (P, R), Péter Niedermüller (NP, L), Tibor Navracsic (NP, R), Zoltan Kovacs (P, R)

**Italy**

Alessandro Di Battista (P, L), Beppe Grillo (P, L), Carlo Calenda (NP, L), Daniela Santanché (P, R), Emma Bonino (NP, L), Enrico Letta (NP, L), Francesco Rutelli (NP, L), Giorgia Meloni (P, R), Giuseppe Conte (P, L), Luigi Di Maio (P, L), Mario Monti (NP, R), Matteo Renzi (NP, L), Matteo Salvini (P, R), Paolo Gentiloni (NP, L), Pier Ferdinando Casini (NP, R), Pier Luigi Bersani (NP, L), Pietro Grasso (NP, L), Roberto Maroni (P, R), Silvio Berlusconi (P, R), Walter Veltroni (NP, L)

**Mexico**

Alberto Anaya Gutiérrez (NP, L), Alejandro Moreno (NP, R), Alfonso Ramírez Cuéllar (P, L), Andrés Manuel López Obrador (P, L), Beatriz Mojica Morga (NP, L), Carlos Puente Salas (NP, L), Carolina Viggiano (NP, R), Clemente Castañeda Hoeflich (NP, L), Dante Delgado (NP, R), Enrique Peña Nieto (NP, R), Felipe Calderón (NP, R), Gabriel Quadri de la Torre (NP, R), Hugo Eric Flores Cervantes (P, L), Héctor Larios (NP, R), Jaime Rodríguez Calderón (P, L), Josefina Vázquez Mota (NP, R), José Antonio Meade (NP, L), Luis Castro Obregón (NP, L), Manuel Granados Covarrubias (NP, L), Marko Cortés (NP, R), Martí Batres (P, L), Patricia Mercado Castro (NP, L), Reginaldo Sandoval Flores (P, L), Ricardo Anaya (NP, R), Roberto Campa Cifrián (NP, L), Yeidckol Polevnsky (P, L)

**Netherlands**

Alexander Pechtold (NP, L), André Rouvoet (NP, L), Arie Slob (NP, R), Diederik Samsom (NP, L), Emile Roemer (P, L), Geert Wilders (P, R), Gert-Jan Segers (NP, R), Henk Krol (NP, L), Jan Peter Balkenende (NP, R), Jesse Klaver (NP, L), Job Cohen (NP, L), Kees van der Staaij (NP, R), Lodewijk Asscher (NP, L), Marianne Thieme (NP, L), Mark Rutte (NP, R), Paul Rosenmöller (NP, L), Sybrand van Haersma Buma (NP, R), Thierry Baudet (P, R), Thom de Graaf (NP, L), Tunahan Kuzu (P, L)

**Norway**

Audun Lysbakken (NP, L), Bjørnar Moxnes (NP, L), Dagfinn Høybråten (NP, R), Erna Solberg (NP, R), Hanna E. Marcussen (NP, L), Jens Stoltenberg (NP, L), Jonas Gahr Støre (NP, L), Knut Arild Hareide (NP, L), Kristin Halvorsen (NP, L), Liv Signe Navarsete (P, L), Rasmus Hansson (NP, L), Thorbjørn Jagland (NP, L), Trine Skei Grande (NP, R), Une Aina Bastholm (NP, L), Åslaug Haga (P, L)

**Poland**

Barbara Nowacka (NP, L), Beata Maria Kusińska (P, R), Donald Tusk (NP, L), Ewa Kopacz (NP, L), Grzegorz Napieralski (NP, L), Janusz Korwin-Mikke (P, R), Janusz Palikot (P, R), Janusz Piechociński (NP, L), Leszek Miller (NP, L), Mateusz Morawiecki (P, R), Małgorzata Maria Kidawa-Błońska (NP, L), Paweł Piotr Kukiz (P, L), Roman Giertych (NP, R), Ryszard Petru (NP, R), Wojciech Olejniczak (NP, L), Władysław Marcin Kosiniak-Kamysz (NP, L), Włodzimierz Czarzasty (NP, L)

**Portugal**

André Ventura (P, R), António Costa (NP, L), Assunção Cristas (NP, R), Carlos Guimarães Pinto (NP, R), Catarina Martins (P, L), José Manuel Barroso (NP, L), José Sócrates (NP, L), Mário Centeno (NP, L), Pedro Passos Coelho (NP, R), Rui Rio (NP, R)

**Slovenia**

Alenka Bratušek (NP, L), Borut Pahor (NP, L), Dejan Židan (NP, L), Franc Bogovič (NP, R), Gregor Golobič (NP, L), Gregor Virant (NP, L), Janez Janša (P, R), Janez Podobnik (NP, R), Karl Erjavec (NP, L), Katarina Kresal (NP, L), Ljudmila Novak (NP, R), Luka Mesec (P, L), Marjan Šarec (P, L), Matej Tonin (NP, R), Miro Cerar (NP, L), Zmago Jelinčič Plemeniti (P, R), Zoran Janković (NP, R)

**Spain**

Albert Rivera (NP, R), Cayo Lara (NP, L), Francesc Homs (NP, L), Gabriel Rufián (P, L), Gaspar Llamazares (NP, L), Joan Ridao (P, L), Josu Erkoreka (NP, L), Mariano Rajoy Brey (NP, R), Oriol Junqueras (P, L), Pablo Casado (NP, R), Pablo Iglesias (P, L), Pedro Sánchez (NP, L), Rosa Díez (NP, L), Santiago Abascal (P, R)

**Sweden**

Alf Svensson (NP, L), Annie Lööf (NP, R), Ebba Busch Thor (P, R), Gudrun Schyman (NP, L), Göran Hägglund (NP, R), Göran Persson (NP, L), Isabella Lövin (NP, L), Jan Björklund (NP, R), Jimmie Åkesson (P, R), Jonas Sjöstedt (NP, L), Lars Ohly (NP, L), Maria Wetterstrand (NP, L), Stefan Löfven (NP, L), Åsa Romson (NP, L)

**United Kingdom**

Arlene Foster (NP, R), Boris Johnson (P, R), Charles Kennedy (NP, L), David Cameron (NP, R), Ed Miliband (NP, L), Gordon Brown (NP, L), Jeremy Corbyn (P, L), Jo Swinson (NP, L), John Swinney (NP, L), Nicola Sturgeon (NP, L), Nigel Farage (P, R), Paddy Ashdown (NP, L), Theresa May (NP, R), Tim Farron (NP, L), William Hague (NP, R)

**United States**[3]

Alexandria Ocasio-Cortez (P, L), Barack Obama (NP, L), Bernie Sanders (P, L), Bill Frist (NP, R), Chuck Schumer (NP, L), Elizabeth Warren (P, L), George Bush (NP, R), Harry Reid (NP, L), Hillary Clinton (NP, L), Joe Biden (NP, L), John Boehner (NP, R), John Kerry (NP, L), John McCain (NP, R), José E. Serrano (NP, L), Kevin McCarthy (NP, R), Mike Pence (NP, R), Mitch McConnell (NP, R), Mitt Romney (NP, R), Nancy Pelosi (NP, L), Paul Ryan (NP, R), Steny Hoyer (NP, L)

## F.7 Bigrams selection

Figure F2 restricts the visualization to bigrams that appear in at least two quarters (left panel) and that occur at least ten times in at least one quarter (right panel). The Figure guides us in our choice of bigram-selection thresholds. The benchmark specification uses bigrams used in at least 10 quarters with a minimum of 10 occurrences in at least one quarter. For robustness checks we also focus on bigrams that were used in at least 5/10 quarters and at least 10/20/25 times in at least one of the quarters.

[3]Donald Trump is not in the US list as his Twitter account (together with that of his digital director, Gary Coby) was permanently suspended after his second impeachment.

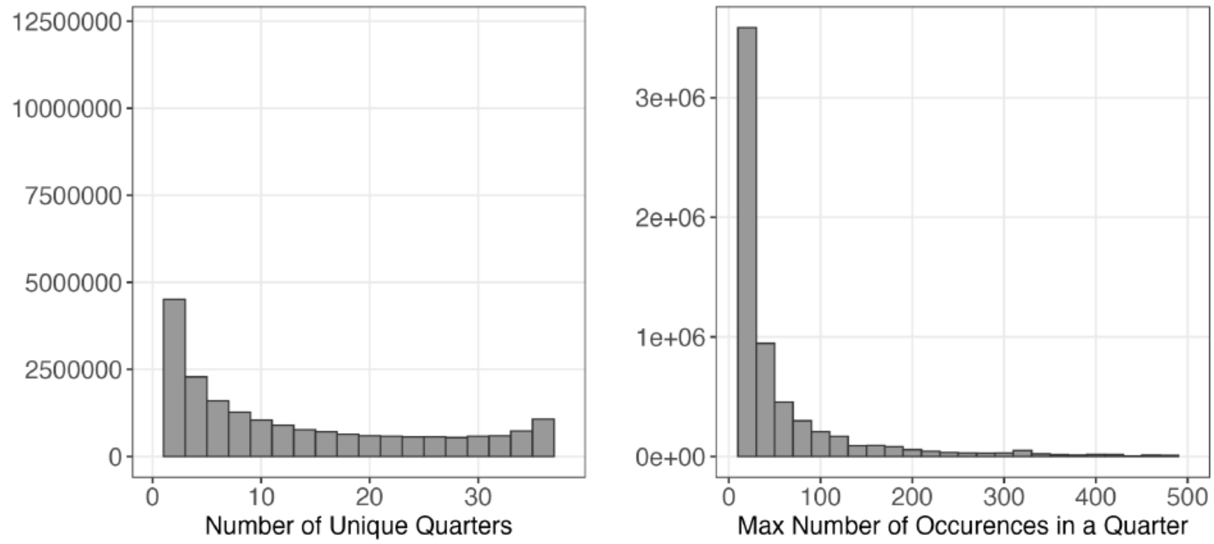



Figure F2: Histograms for the two measures of bigram frequency: No. of unique quarters a bigram was used (left panel), Max no. of occurrences of a bigram in a quarter (right panel). Low values of both measures are omitted.

## F.8 Topics Classification Prompt

To create the training sample using ChatGPT the following prompt was used:

"You are a helpful research assistant.
I will give you a tweet from a politician around the world (translated into English). Classify it into one or more specific topics or return "no topic" if none apply. Here are the topics to classify into, along with descriptions. Some topics also include a list of keywords: these are examples of common expressions used for that topic, not strict rules. Use them as hints, but still classify semantically when the topic is implied even without those words. Use only the labels shown in parentheses. Each tweet must match exactly one or more of these labels.

When classifying, return only the labels:
**Immigration and integration**: topics related to the movement of people across borders, refugee issues, their integration into host countries, acquisition of citizenship rights, treatment of illegal migrants.
**Environmental policy and climate change**: topics related to the natural environment, pollution, long-term changes in global weather patterns and their environmental impact, and policies addressing climate change and environmental impact.
**Foreign policy and defence**: topics related to international relations, bilateral or multilateral agreements, military policies, defence spending.
**Rights**: topics related to gender, civil, human, or political rights. Including but not limited to

gender equality, women's rights, abortion, LGBTQ+ issues, diversity equity and inclusion policies, individual freedoms and rights (including free speech and personal liberties), political freedoms (such as voting, assembly, freedom of the press, and any kinds of political participation), bioethical issues.

**Security, crime, and violence**: topics related to internal security (e.g. "We need more police patrolling", "There is a security threat"), policing, crime, guns license, acts of violence (e.g. news of a murder).

**Public health and healthcare policy**: topics related to public health policies, health crises, vaccines, healthcare systems.

**European Union affairs**: topics specifically related to the European Union, including its institutions and policies.

**Industrial policy, business, finance**: topics related to regulation, productivity, banks, credit, finance, market competition, antitrust, business policies, small and medium enterprises, industrial subsidies, and topics mentioning any private firm or corporate entity.

**International trade and globalization**: topics related to international trade, tariffs, globalization, imports, commercial wars, unfair competition from China.

**Energy and resources**: topics related to energy sources, such as electricity, hydrogen, solar, wind power, nuclear, coal, oil, and natural gas.

**Inequality**: topics related to disparities in wealth, income, or socio-economic status, poverty, redistribution, policies providing social assistance to poor people.

**Fiscal policy and taxation**: topics related to taxes, government revenues, budgets, public debt and economic policy related to taxation, government spending, pension systems.

**Inflation and monetary policy**: topics related to rising price levels, cost of living, and to the related monetary policy measures and instruments (e.g. interest rates, quantitative easing/tightening, central banks policies).

**Labour market and workers' rights**: topics related to employment, unemployment, work injuries, job markets, labour rights, labour market policies, and labour market institutions (such as minimum wages, unions, employment protection legislation).

**Connecting**: topics related to non-political but personal communication such as greetings, weather or holidays. Keywords: greetings, holiday, birthday.

**Elections and voting**: topics related to elections, voting or political campaigns. Keywords: election, vote, campaign, candidate.

**Self-promotion**: when politicians promote themselves or their work e.g. "I'm on air today", "Check out my interview", "We have accomplished a lot during our term" etc. Keywords: join me, live, watch, interview.

**Anti-establishment**: Covers tweets which are

(a) anti-system: idea that the current political system is broken, corrupt, rigged and undemocratic.

(b) anti-elites: idea that elites are irresponsible, dishonest, unresponsive to the needs of the people,

guilty of cronyism and only look after their own interests. By elites we mean politicians, bureaucrats and civil servants (including international organizations like EU, IMF, WTO), economic elites (eg. big business, millionaires), central banks, top universities and establishment media.
(c) emphasizing the conflict between "the people" against "the elites/establishment": idea that society is split into two antagonistic groups: "the people" and "the corrupt elite".
Do not tag anti-establishment when:
(i) The target is a single politician or a specific party without generalizing to political class / elites / establishment.
(ii) The target is a foreign leader or regime (use foreign policy and possibly rights/security).
(iii) It is ordinary policy disagreement without elites/system are illegitimate framing.

Rules:
Return only the labels as a comma-separated list (e.g., "labour market, taxes"). If no topic applies, return exactly "no topic". A tweet can belong to multiple topics. If multiple topics apply, list all of them. If the tweet does not fit any of these topics, output only "no topic". Be concise and accurate in matching the topics."

## F.9 Back Translation

To assess the quality of English translations, a sample of tweets per country was translated back into their original languages. For countries where many tweets were originally written in English and therefore excluded from the validation sample, larger samples were initially drawn to maintain representativeness. Before being compared, punctuation and language-specific stopwords were removed, and all texts have been lowercased. Tweets have been translated using Google Translate. The similarity between each back-translated tweet and its original version was then evaluated using multiple text similarity metrics.

First, we compute Jaccard Similarity, which measures how many words the two tweets have in common relative to the total number of unique words across both texts:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|},$$

where $A$ is the set of words in the reference text and $B$ is the set of words in the translated text. Second, we compute the BLEU Score: define unigram (bigram) precision as the number of matching unigrams (bigrams) over the number of unigrams (bigrams) in the candidate text. The score is based on modified n-gram precision, where matching n-gram counts are clipped by their maximum frequency in the reference text in order to avoid rewarding repeated words. Let $p_1$ and

$p_2$ denote the modified unigram and bigram precision, the BLEU score is computed as:

$$\text{BLEU} = BP \cdot \exp\left(\frac{1}{2}\log p_1 + \frac{1}{2}\log p_2\right) = BP \cdot \sqrt{p_1 p_2},$$

where $BP$ is the brevity penalty, whose goal is to discourage overly short outputs, and is defined as:

$$BP = \begin{cases} 1 & \text{if } c > r, \\ \exp(1 - r/c) & \text{if } c \leq r, \end{cases}$$

with $c$ denoting the candidate length and $r$ the reference length. Third, we compute Cosine Similarity, which measures the semantic similarity between two texts by comparing their vector representations (embeddings) generated using the multilingual SentenceTransformer model `paraphrase-multilingual-MiniLM-L12-v2`, which supports more than 50 languages. Given the two embedding vectors $A$ and $B$ of the reference and translated texts, cosine similarity is computed as

$$\text{Cosine Similarity} = \frac{A \cdot B}{\|A\|\|B\|}.$$

Table F5: Back Translation Metrics

| Jaccard | BLEU (Only Bigrams) | BLEU | Cosine Similarity | Tweets | Language |
|---|---|---|---|---|---|
| 0.65 | 0.59 | 0.65 | 0.96 | 106 | Spanish |
| 0.62 | 0.55 | 0.63 | 0.97 | 100 | Italian |
| 0.61 | 0.55 | 0.62 | 0.96 | 98 | Swedish |
| 0.56 | 0.45 | 0.54 | 0.96 | 96 | German |
| 0.54 | 0.46 | 0.54 | 0.94 | 95 | Czech |
| 0.65 | 0.59 | 0.65 | 0.96 | 88 | Portuguese |
| 0.56 | 0.51 | 0.58 | 0.95 | 83 | Danish |
| 0.61 | 0.53 | 0.61 | 0.95 | 83 | Slovene |
| 0.60 | 0.55 | 0.62 | 0.95 | 75 | Norwegian |
| 0.54 | 0.45 | 0.53 | 0.95 | 66 | Dutch |
| 0.40 | 0.31 | 0.40 | 0.92 | 61 | Finnish |
| 0.68 | 0.61 | 0.68 | 0.96 | 61 | French |
| 0.59 | 0.52 | 0.60 | 0.96 | 51 | Polish |
| 0.48 | 0.35 | 0.44 | 0.90 | 48 | Hungarian |

# G Monte Carlo simulation for number of speakers and unobserved groups

Figure 3 raises two concerns: (1) partisanship for the random group exceeds 0.5, (2) partisanship for the random group of politicians exhibits pre-electoral divergence (qualitatively similar to the dynamics for the populist divide). Here we conduct a Monte Carlo simulation to examine whether these patterns could be due to the small number of politicians in the sample (relative to the number of possibly relevant political types).

We find that the observed patterns for the random groups are likely to be driven by both the small number of politicians and the presence of unobserved political types. However, it is important to note that the results obtained in this section reflect the properties of the penalized estimators when being used on different samples, not the properties of the samples. We verify that this is indeed the case by estimating random groups partisanship with the leave-out estimator, which consistently yields partisanship of 0.5 even in the presence of unobserved groups.

We keep the same set of bigrams used in the benchmark specification, but generate random bigram usage along with random samples of politicians for 20 sessions. For the politician samples, we construct a populist dummy, preserving the observed share of populist politicians (approximately 0.225), and vary the number of politicians per sample across $N \in \{50, 250, 500\}$. The median number of active politicians in our sample is 287, so the sample of 250 politicians is the closest to our data configuration. For the bigrams sample, we randomly draw bigram usage using Dirichlet sampling, keeping verbosity close to its observed value. We try to keep the distribution of bigrams such that partisanship matches the one estimated using the real data. Having obtained the simulated sample of bigrams used by each politician: (i) we estimate the parameters of the utility of speech; (ii) compute $\hat{q}_{ijt}^{P_i}$, for each politician $i$ bigram $j$ and session $t$; (iii) obtain individual - session specific partisanship measure, $\pi_{it}$; (iv) compute average partisanship, $\pi_t$, over the entire sample of politicians.

The bigrams sample is simulated under two scenarios, which we now describe, to explore the importance of a small number of politicians and of unobserved political types.

**Scenario I.** Here we explore the consequences of changing the number of politicians in the sample. First, we randomly select 1,000 bigrams and classify them as populist. When generating bigram usage, we boost the frequency of these bigrams for *populist* politicians by a factor of two. Second, among the 20 sessions we simulate, we treat the first five as pre-electoral, i.e., sessions in which the effect of political types on bigram choice (parameter $\phi$) is higher. To capture this, we apply an additional boost to the same set of 1,000 bigrams in the first five sessions, of about 20% on average. In other words, populist politicians consistently use this set of 1,000 bigrams more frequently across all 20 sessions, but especially so during the first five.

The results of this simulation are presented in Figure G1. The solid lines show the average partisanship outcome, $\pi_t$, for the populist versus non-populist division, while the dotted lines depict the average $\pi_t$ for ten random binary partitions that do not correspond to the populist–non-populist split. Vertical segments are 95% confidence intervals.

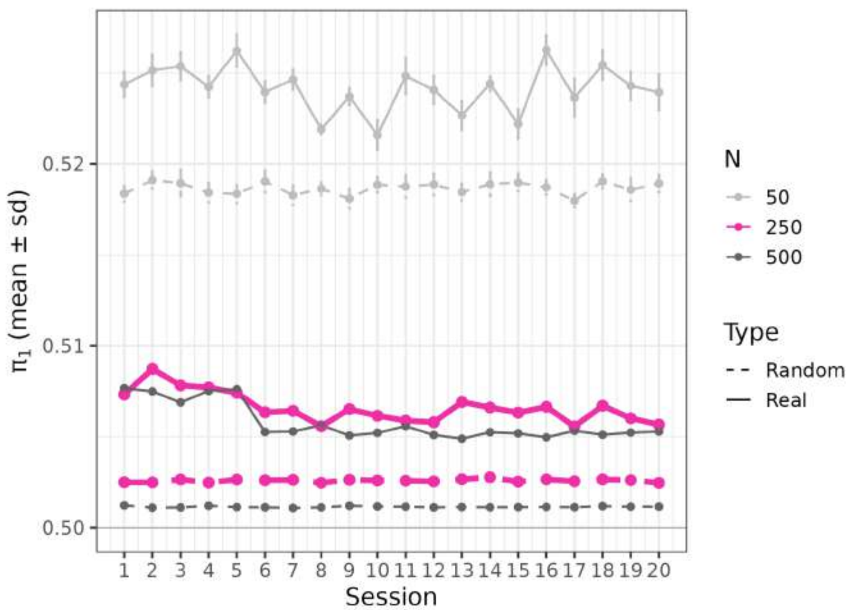



Figure G1: Average partisanship estimates over politicians and sessions for real and random populist groups. No unobserved electorally relevant groups.

As the number of politicians increases, populist partisanship is not affected once it reaches 250 politicians, but partisanship of the random groups (the dotted lines) converges toward 0.5. This addresses concern (1), suggesting that the relatively large observed values of partisanships for the random groups are partly driven by the small number of politicians in our sample.

Nevertheless, in these simulations partisanship of the random groups is not sensitive to varying parameter $\phi$, since it has the same pattern in the first five sessions as in the remaining ones (the difference between the first sessions and the rest is statistically significant for populist partisanship, but not for partisanship of the random groups). This suggests that the number of politicians, on its own, cannot explain the observed pre-electoral pattern of the random group partisanship in Figure 4.

**Scenario II.** To explain this pre-electoral pattern, in Figure G2 we explore the consequences of also having unobserved political types. Here we keep the same data generating process, except that in addition to populist and non-populist, we generate two additional political types within the populist and non-populist subsets, boosting for each type the usage of a specific set of bigrams. Moreover, similar to the additional boost in sessions 1-5 for the populist bigrams, we boost the usage of these specific sets of bigrams in sessions 1-5 too, thus making these four political groups electorally relevant. We then compute average partisanship between two groups of politicians: populist vs non-populist (the solid lines) and an entirely random partition (the dotted lines). Again we repeat these simulations for three sample sizes of politicians, 50, 250 and 500. The results of these simulations are depicted in Figure G2.

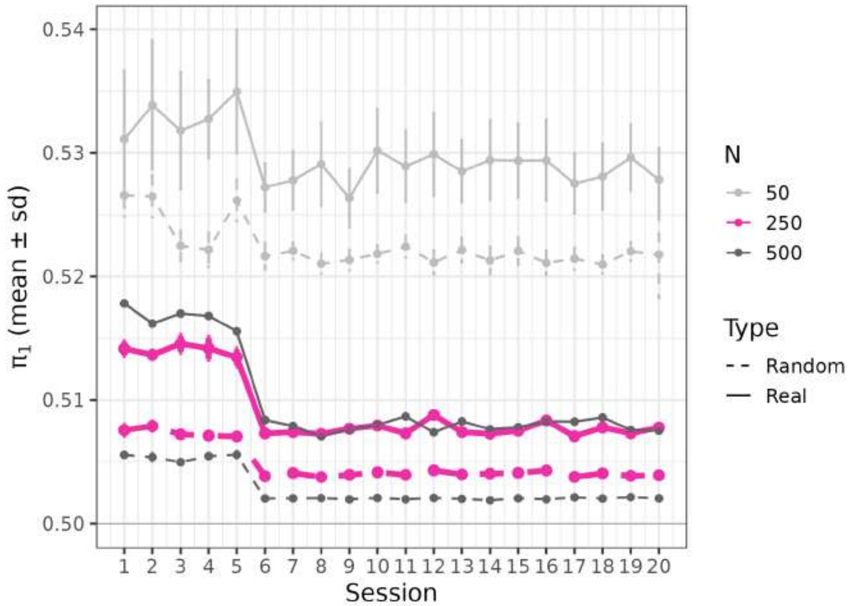



Figure G2: Average partisanship estimates over politicians and sessions for real and random populist groups. Four unobserved electorally relevant groups.

Partisanship in the random groups now is indeed much larger in Sessions 1-5, where unobserved political types have a larger effect on bigram usage, than in the remaining sessions, and this difference is statistically significant. This pattern suggests that some political types are not fully accounted for in our framework. Note that the sample size of politicians continues to be relevant also in these simulations.

This suggests that what matters is not merely the number of politicians, but the number of politicians relative to the number of relevant political types. If the number of politicians is not sufficiently large relative to the number of relevant political types, the two random groups differ systematically in their bigram usage and this is captured by the partisanship statistics.

# H Tables and Figures

## H.1 Tables

Table H1: Politicians by ideology and populism (counts with row percentages)

| | Non-populist | Populist | Total |
|---|---|---|---|
| Left | 185 (83.0%) | 38 (17.0%) | 223 (60.8%) |
| Right | 98 (68.1%) | 46 (31.9%) | 144 (39.2%) |
| Total | 283 (77.1%) | 84 (22.9%) | 367 (100%) |

Table H2: Wald Tests: Coefficient differences from $\beta_{\text{quarter diff=0}}$

| Hypothesis | F-statistic | p-value | Dep. Variable |
|---|---|---|---|
| $\beta_{\text{quarter diff=1}} = \beta_{\text{quarter diff=0}}$ | 3.775 | 0.0521 | $\pi_{it}$ |
| $\beta_{\text{quarter diff=2}} = \beta_{\text{quarter diff=0}}$ | 4.479 | 0.0343 | $\pi_{it}$ |
| $\beta_{\text{quarter diff=3}} = \beta_{\text{quarter diff=0}}$ | 1.076 | 0.3000 | $\pi_{it}$ |

Table H3: Unique Elections by Quarter

| | | | | | |
|---|---|---|---|---|---|
| No. of unique countries per election quarter | 1 | 2 | 3 | 4 | 5 |
| No. of occurrences | 12 | 10 | 3 | 2 | 2 |

*Notes.* Only calendar quarters used in the main analysis are considered. Elections that fall outside of the study window are not considered.

Table H4: BERT vs Manual Labeling Metrics

| | Accuracy | F1 Score |
|---|---|---|
| ***Policy Topics*** | | |
| Business | 0.91 | 0.64 |
| Climate | 0.96 | 0.83 |
| Energy | 0.97 | 0.87 |
| European Union | 0.98 | 0.89 |
| Foreign Policy | 0.92 | 0.61 |
| Health | 0.97 | 0.86 |
| Immigration | 0.98 | 0.88 |
| Inequality | 0.92 | 0.55 |
| Inflation | 0.96 | 0.55 |
| Labor Market | 0.96 | 0.83 |
| Rights | 0.93 | 0.68 |
| Security | 0.94 | 0.69 |
| Taxes | 0.97 | 0.86 |
| Trade | 0.96 | 0.42 |
| ***Non-Policy Topics*** | | |
| Anti Establishment | 0.94 | 0.49 |
| Personal Communication | 0.96 | 0.53 |
| Elections | 0.96 | 0.67 |
| Self Promotion | 0.81 | 0.33 |

*Notes:* Accuracy is defined as the number of correct predictions divided by the total number of predictions. The F1 Score is the harmonic mean of precision and recall: it is twice their product divided by their sum. Precision is the number of true positives divided by the number of all predicted positives, while recall is the number of true positives divided by the number of actual positives.

## H.2 Figures

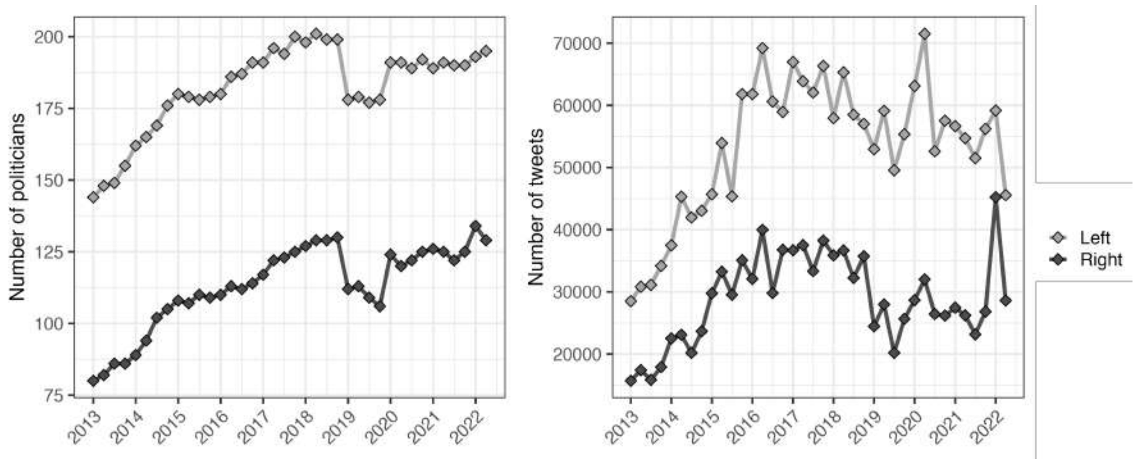



Figure H1: Number of politicians active on Twitter per calendar quarter, and number of tweets, per Left-Right divide

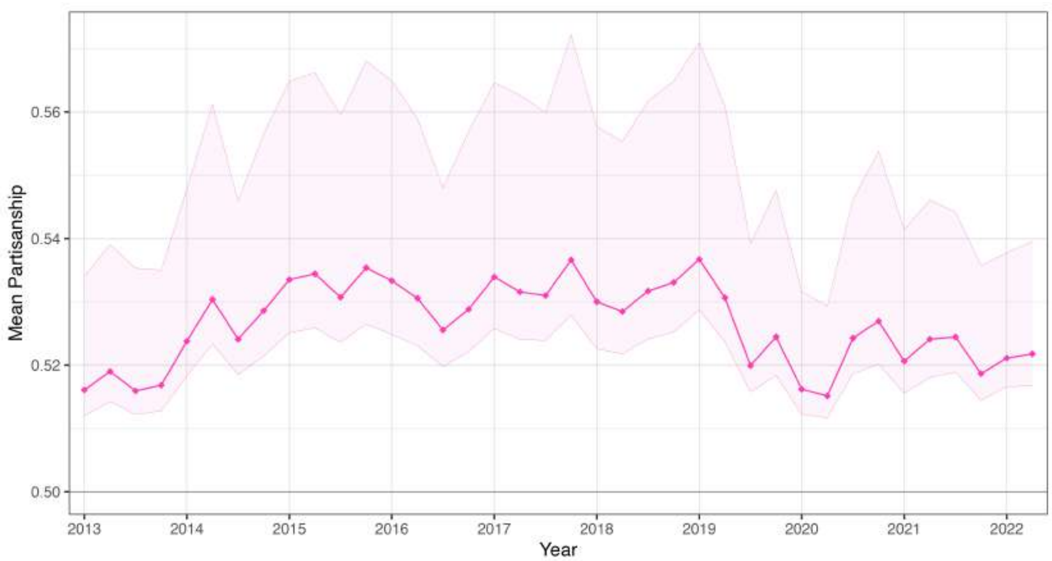



Figure H2: Average partisanship with 90% CI

*Notes.* Figure plots average values and confidence intervals of populist partisanship over the study period. Partisanship is defined as in Equation 11. Average values are defined as in Equation 12. Confidence Intervals are constructed via subsampling, as described in Section D.

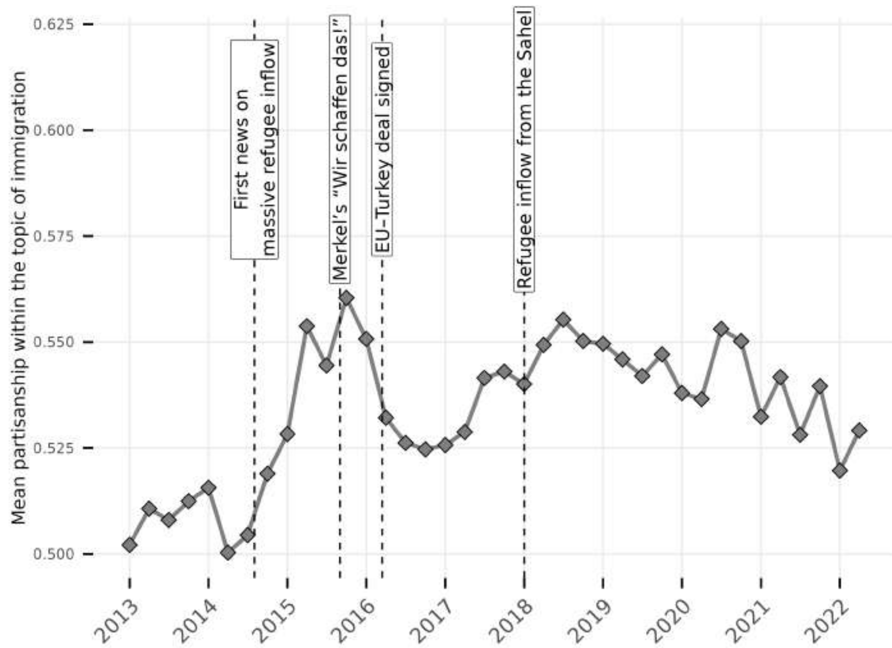



Figure H3: Partisanship within the topic of immigration

*Notes.* The figure reports the average P–NP partisanship by calendar quarter, computed using the sample of bigrams that belong to the immigration topic and restricting the sample to European countries. Partisanship is defined as in Equation 11. Average values are defined as in Equation 12.

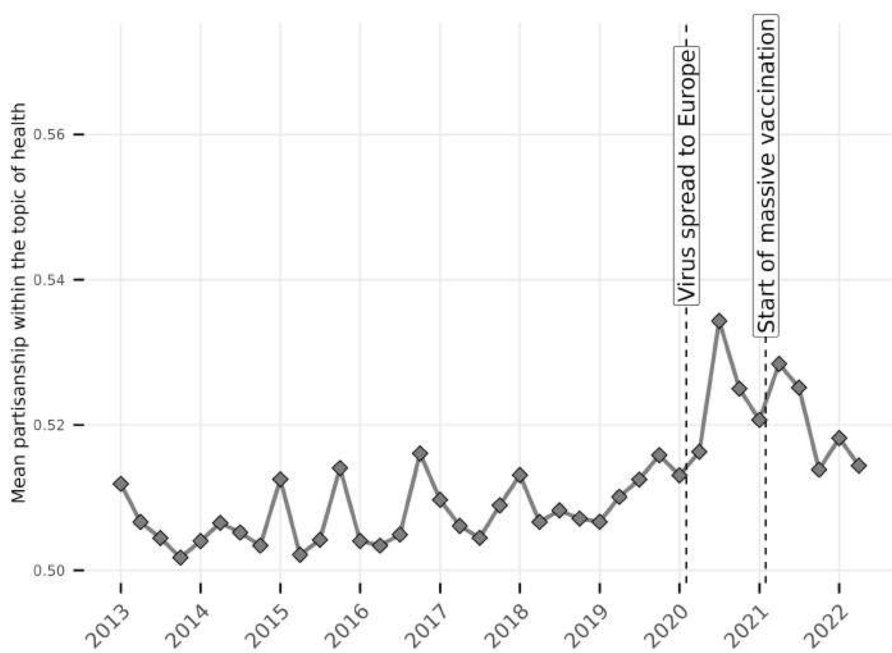



Figure H4: Partisanship within the topic of public health

*Notes.* The figure reports the average P–NP partisanship by calendar quarter, computed using the sample of bigrams that belong to the public health topic. Partisanship is defined as in Equation 11. Average values are defined as in Equation 12.

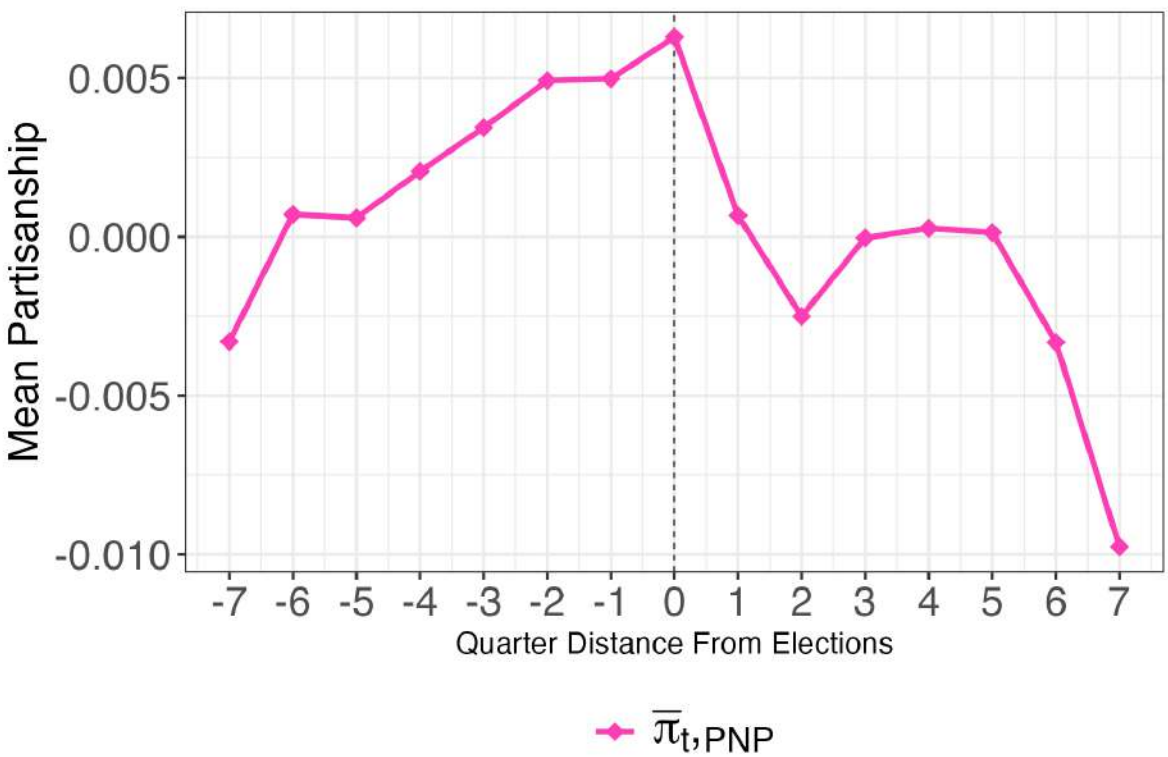
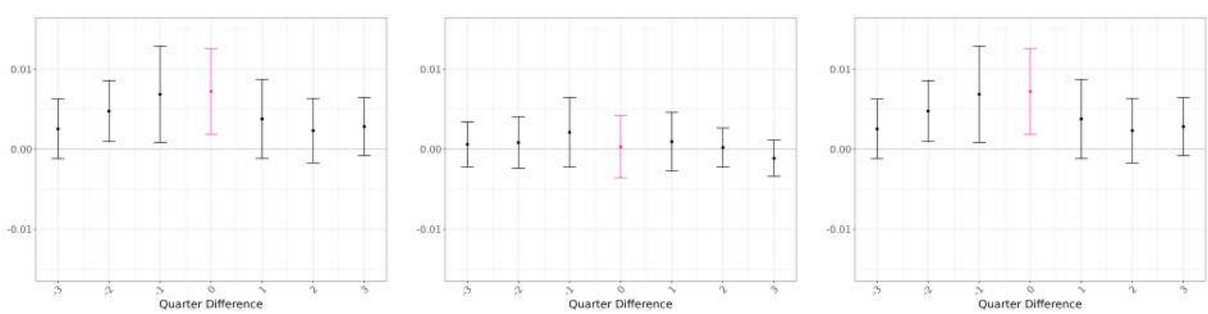



Figure H5: Average estimated $\pi_{it}$ per quarter difference from elections (net of calendar-quarter fixed effects).

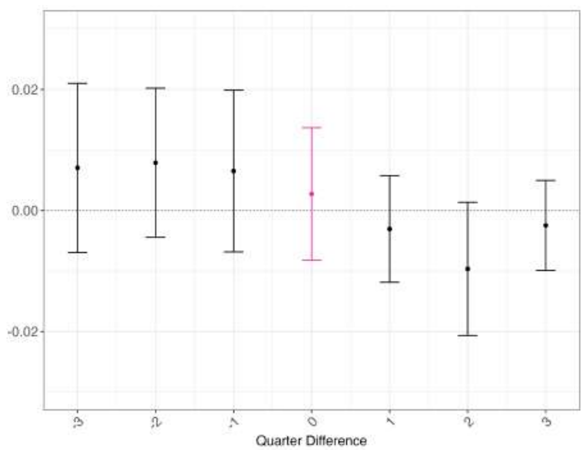
*Notes.* $\bar{\pi}_{t,PNP}$ refers to average populist partisanship net of calendar-quarter fixed effects. Partisanship is defined as in Equation 11. Average values are defined as in Equation 12.



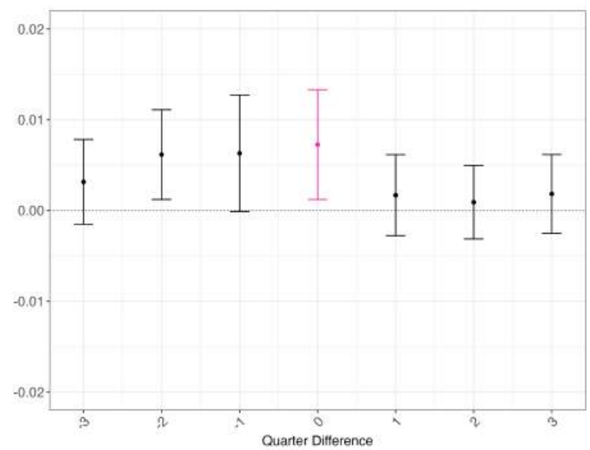

(a) Populist vs Non-Populist (b) Left vs Right (c) Four political types

Figure H6: Event study estimates from the model with four political types (including both populist-non-populist and left-right dimensions)

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13. Outcome variable is populist partisanship (left panel), left-right partisanship (middle panel) and four types partisanship (right panel). Errors are clustered at the elections level, 95% confidence intervals are reported. Partisanship with several political types is defined as in Section C.




(a) English-speaking countries




(b) Not English-speaking countries

Figure H7: Event study estimates for $\pi_{it}$ for countries with English and other languages as main official language

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13 on the subsamples of English-speaking countries (left) and not English-speaking countries (right). Errors are clustered at the elections level, 95% confidence intervals are reported. Partisanship is defined as in Equation 11.

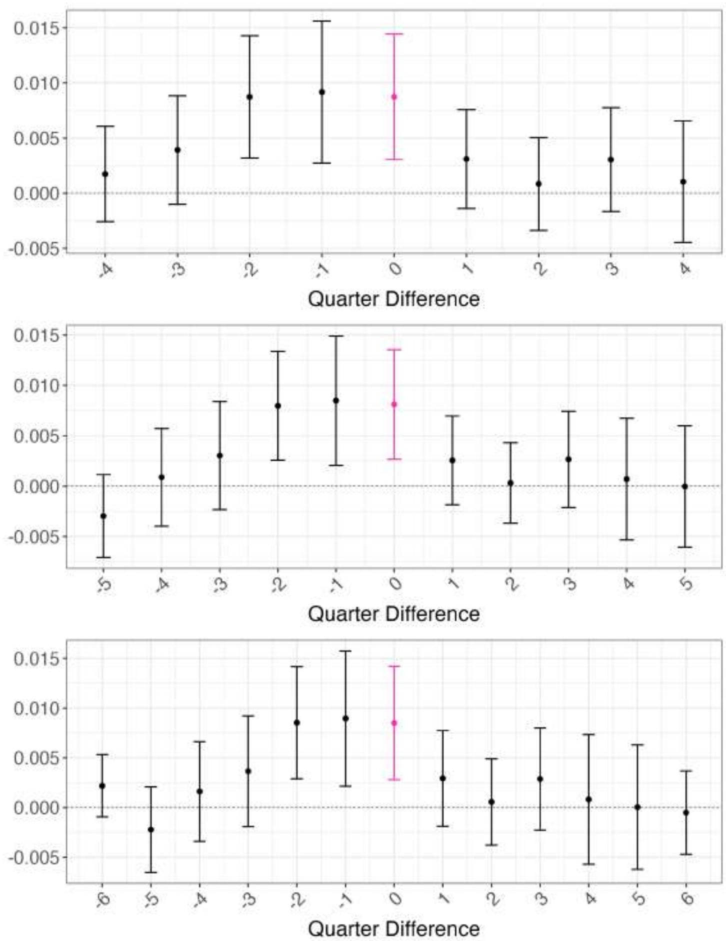



Figure H8: Event study results. Bigram selection threshold is 10-10. Event window is 4 quarters, 5 quarters and 6 quarters.

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13 using different definitions of electoral window – 4 quarters (upper panel), 5 quarters (middle panel), 6 quarters (bottom panel). Errors are clustered at the elections level, 95% confidence intervals are reported. Outcome variable is a populist partisanship, defined as in Equation 11.

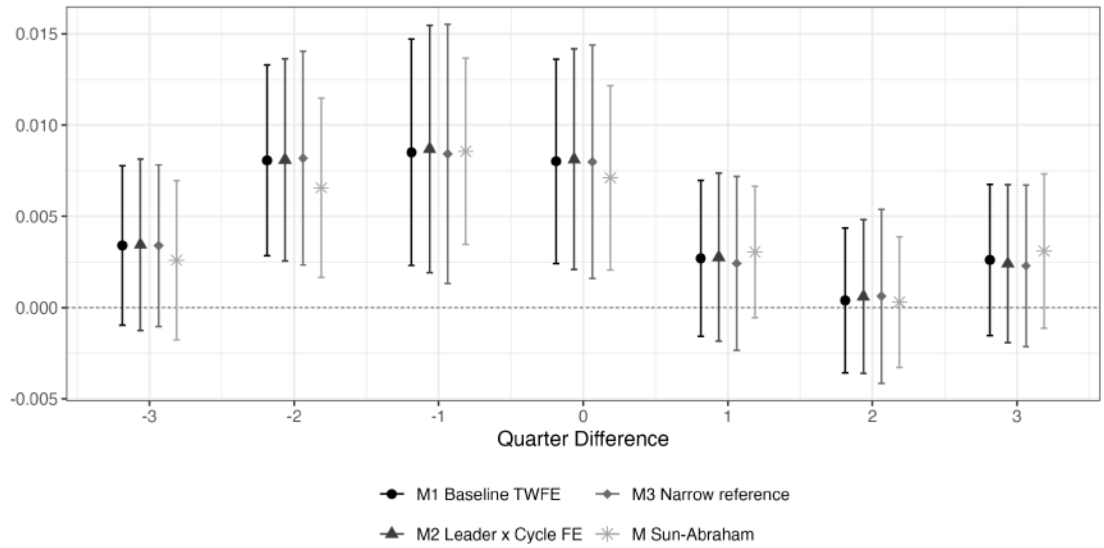



Figure H9: Partisanship event study results, TWFE robustness

*Notes.* Partisanship is defined as in Equation 11. Each panel plots coefficients from four TWFE event-study specifications. M1 is the baseline (leader, quarter, and country × electoral-cycle fixed effects; errors clustered by electoral cycle). M2 replaces the leader FE with a leader × electoral-cycle FE, restricting identification to within-leader within-cycle variation. M3 uses a narrow reference window (quarters 4–6 from the election only). Bars are 95% confidence intervals. M Sun-Abraham plots an interaction-weighted estimator, which aggregates election-cohort-specific coefficients with non-negative weights to guard against bias from heterogeneous effects across elections (Sun and Abraham, 2021).

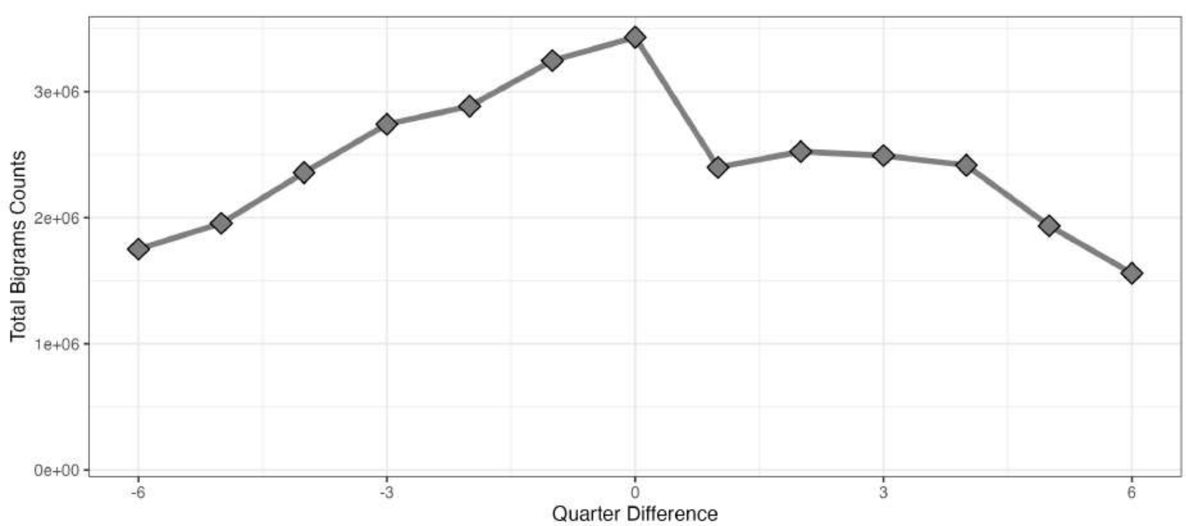



Figure H10: Number of bigrams used per quarter, difference from the elections

*Notes.* This graph plots the total number of all bigrams used in tweets at a given quarter difference from elections.

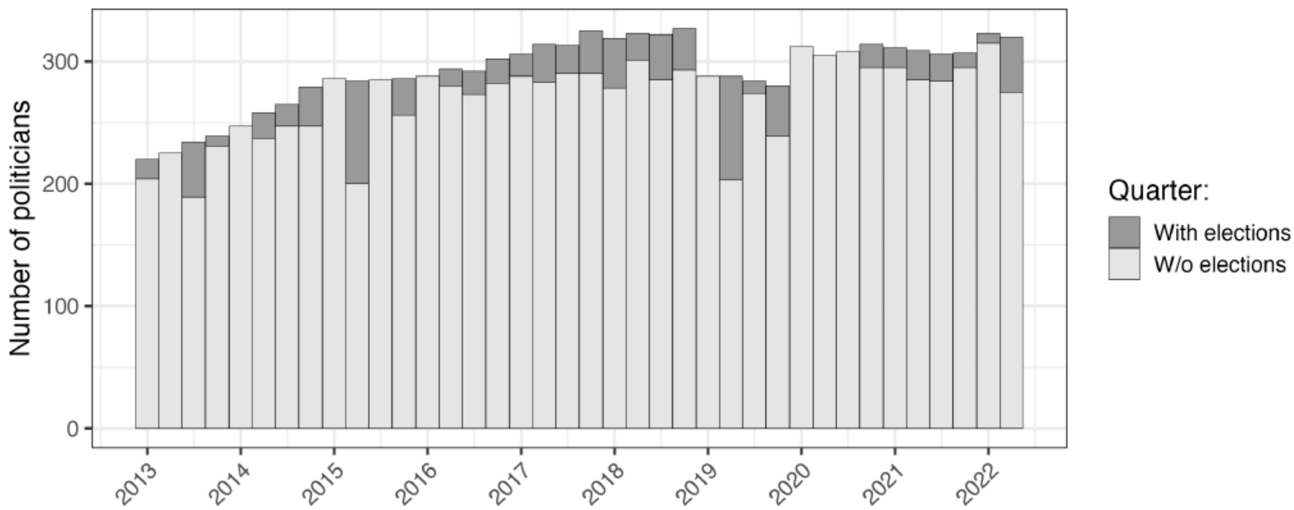



Figure H11: Number of politicians per calendar quarter

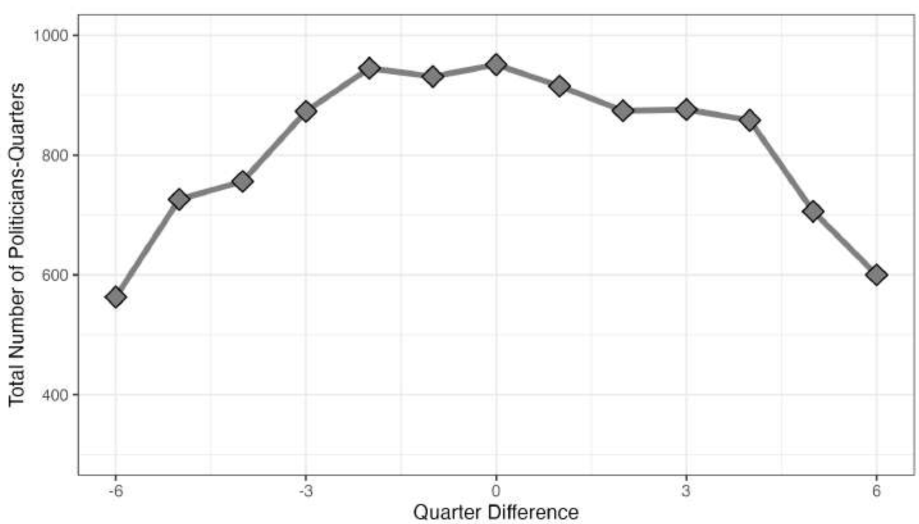



Figure H12: Number of unique quarter-politicians observations across the electoral cycle

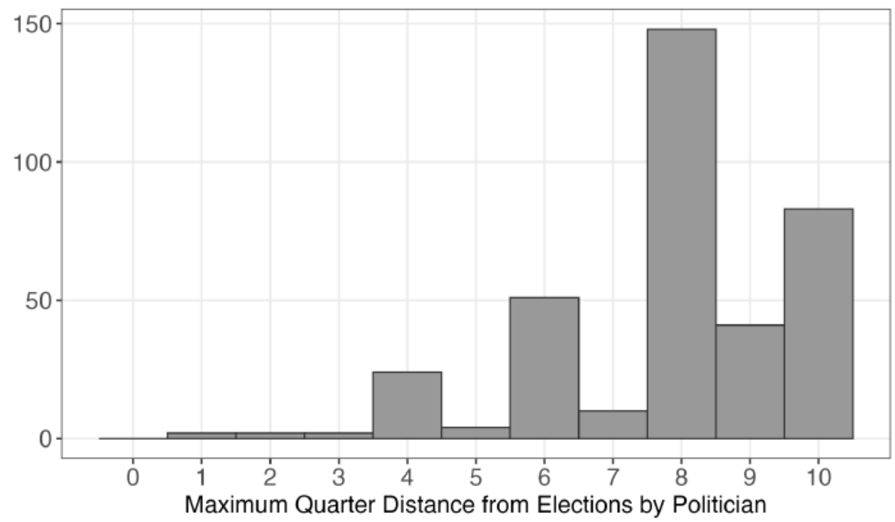



Figure H13: Maximum absolute quarter distance from the elections for politicians

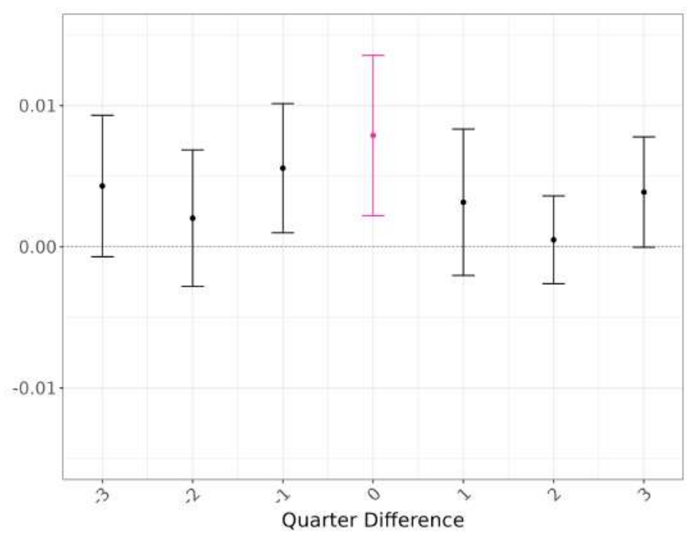
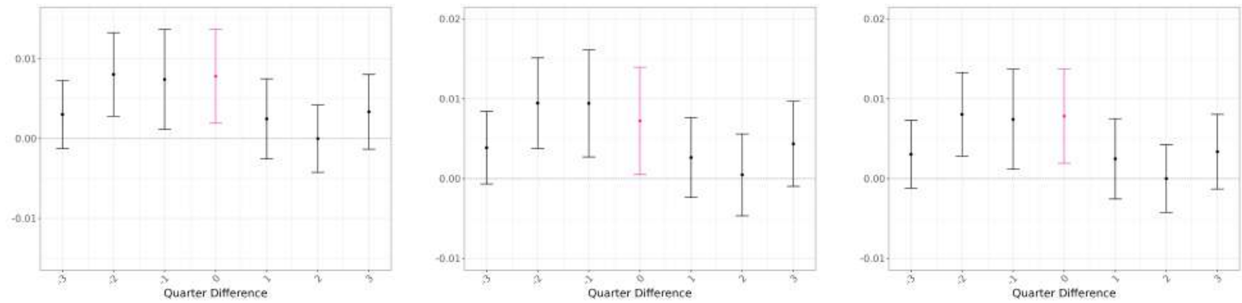



Figure H14: Event study estimates for $\pi_{it}$ based on a restricted random sample of 220 politicians per quarter.

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13 on the restricted sample of 220 politicians randomly selected in each quarter. Errors are clustered at the elections level, 95% confidence intervals are reported. Outcome variable is populist partisanship, defined as in Equation 11.




(a) 5-10 (b) 5-20 (c) 10-25

Figure H15: Event study estimates for $\pi_{it}$ using the 5-10 bigram selection threshold (left panel), 5-20 (center panel) and 10-25 (right panel)

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13 on the samples with different bigram selection thresholds. Left panel includes all bigrams used in at least 5 quarters and at least 10 times in at least one quarter. Center panel includes all bigrams used in at least 5 quarters and at least 20 times in at least one quarter. Right panel includes all bigrams used in at least 10 quarters and at least 25 times in at least one quarter. Errors are clustered at the elections level, 95% confidence intervals are reported. Outcome variable is populist partisanship, defined as in Equation 11.

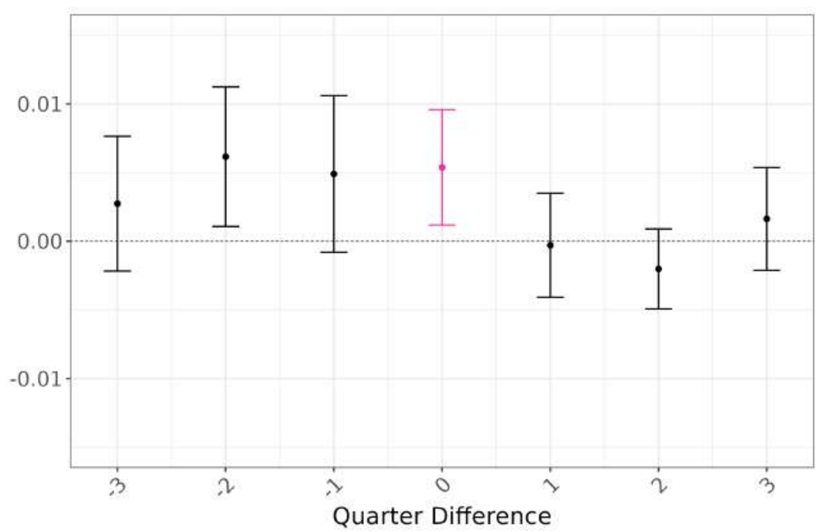



Figure H16: Event study estimates for $\pi_{it}$ with ambiguous politicians coded as non-populists

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13. Outcome variable is populist partisanship, defined as in Equation 11. Partisanship is estimated by assigning ambiguous politicians to the non-populist group (rather than to the populist group, as in the main text). Errors are clustered at the elections level, 95% confidence intervals are reported.

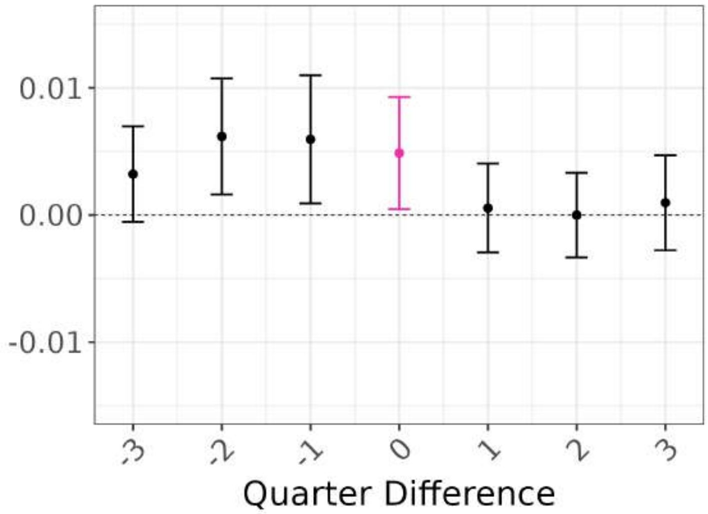



Figure H17: Event study estimates for $\pi_{it}$ on the sample of bigrams without mentions of specific parties and politicians

*Notes.* Figure plots point estimates and confidence intervals for $\beta_{d(i,t)}$ from Specification 13. Outcome variable is populist partisanship, defined as in Equation 11. Partisanship is estimated on a subset of bigrams that excludes bigrams mentioning specific parties or politicians. Errors are clustered at the elections level, 95% confidence intervals are reported.

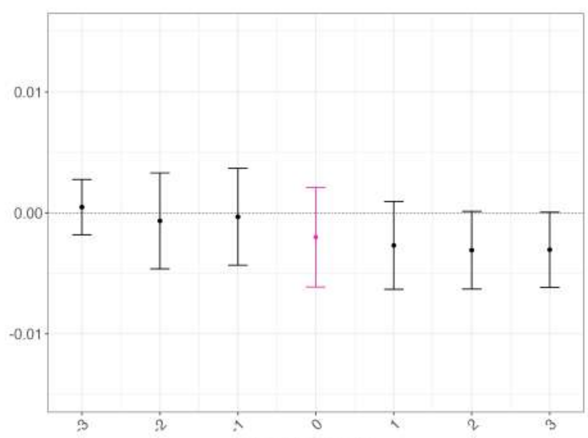



Figure H18: Event studies for $\pi_{it}$, Left-Right partisanship measure

*Notes.* Point estimates and confidence intervals are plotted for $\beta_{d(i,t)}$ from Specification 13. Outcome variable is left-right partisanship, with affiliation defined by the Manifesto RILE variable (instead of the ChatGPT classification). Partisanship is defined as in Equation 11. Errors are clustered at the elections level, 95% confidence intervals are reported.

# References

Chatterjee, Arindam and Soumendra Nath Lahiri. 2011. "Bootstrapping lasso estimators." *Journal of the American Statistical Association* 106(494):608–625.

Colonnelli, Emanuele, Valdemar Pinho Neto and Edoardo Teso. 2025. "Politics at work." *American Economic Review* 115(10):3367–3414.

Sun, Liyang and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of econometrics* 225(2):175–199.