On-Line Appendix for
"Misspecified Politics and the Recurrence of Populism"

(Gilat Levy, Ronny Razin & Alwyn Young - September 2021)

I: Proof of (A28a), (A28b), & (A28c)

We begin by proving (A28a) and (A28b), turning to (A28c) at the end. We start by calculating expressions for $f(c\mathbf{M}^{-1})$ and $f(c^2\mathbf{M}^{-2})$ using (A23) in the text and the Sherman-Morrison formula:

(I.1) $$f(c\mathbf{M}^{-1}) = c\boldsymbol{\beta}_s'\left[\mathbf{V} + \frac{\boldsymbol{\beta}_s\boldsymbol{\beta}_s'}{\boldsymbol{\beta}'\boldsymbol{\beta}}\right]^{-1}\boldsymbol{\beta}_s, \text{ where } \mathbf{V} = \frac{\mathbf{X}_{ss}'\mathbf{X}_{ss}}{t_c R} + \frac{t_s + t_c}{t_c}\frac{\sigma_n^2}{R}\mathbf{I}_{k_s}$$

$$= c\boldsymbol{\beta}_s'\left[\mathbf{V}^{-1} - \frac{\mathbf{V}^{-1}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'\mathbf{V}^{-1}/\boldsymbol{\beta}'\boldsymbol{\beta}}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}}\right]\boldsymbol{\beta}_s \quad \Rightarrow \quad f(c\mathbf{M}^{-1}) = \frac{c\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}}$$

$$f(c^2\mathbf{M}^{-2}) = c^2\boldsymbol{\beta}_s'\left[\mathbf{V} + \frac{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}{\boldsymbol{\beta}'\boldsymbol{\beta}}\right]^{-1}\left[\mathbf{V} + \frac{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}{\boldsymbol{\beta}'\boldsymbol{\beta}}\right]^{-1}\boldsymbol{\beta}_s$$

$$= c^2\boldsymbol{\beta}_s'\left[\mathbf{V}^{-1} - \frac{\mathbf{V}^{-1}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'\mathbf{V}^{-1}/\boldsymbol{\beta}'\boldsymbol{\beta}}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}}\right]\left[\mathbf{V}^{-1} - \frac{\mathbf{V}^{-1}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'\mathbf{V}^{-1}/\boldsymbol{\beta}'\boldsymbol{\beta}}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}}\right]\boldsymbol{\beta}_s$$

$$\Rightarrow \quad f(c^2\mathbf{M}^{-2}) = \frac{c^2\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s}{(1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta})^2} \quad \Rightarrow \quad \frac{f(c^2\mathbf{M}^{-2})}{f(c\mathbf{M}^{-1})^2} = \frac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s}{(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2}$$

We then use the spectral decomposition of $\mathbf{V}$ to create two key expressions:

(I.2) $$\text{(a) } \boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s = \sum_{i=1}^{k_s}\lambda_i a_i^2 \quad \& \quad \text{(b) } \frac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s}{(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2} = \frac{\sum_{i=1}^{k_s}\lambda_i^2 a_i^2}{\sum_{i=1}^{k_s}\lambda_i a_i^2\sum_{i=1}^{k_s}\lambda_i a_i^2},$$

where $\lambda_1 \geq ... \geq \lambda_i \geq ... \geq \lambda_{k_s}$ are the ordered eigenvalues of $\mathbf{V}^{-1}$ and the a_i the inner-products of the associated eigenvectors with $\boldsymbol{\beta}_s$, i.e., $\boldsymbol{a} = \mathbf{E}'\boldsymbol{\beta}_s$. As noted earlier, the eigenvalues of $\mathbf{V}^{-1}$ are the inverse of those of $\mathbf{V}$, while adding c times the identity matrix to a matrix increases all of its eigenvalues by c, so we know that:

(I.3) $$\lambda_i = \frac{t_c R}{\gamma_i + (t_s + t_c)\sigma_n^2}$$

where $\gamma_1 \leq ... \leq \gamma_i \leq ... \leq \gamma_{k_s}$ are the ordered eigenvalues of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$. While the λ_i are in descending order, the corresponding γ_i are in ascending order, as the two are inversely related. We note that the eigenvalues of $\mathbf{V}$ are bounded between $(1+t_s/t_c)\sigma_n^2/R$ and $(t_s/t_c) + k_s(1+t_s/t_c)\sigma_n^2/R$,[1] so in manipulating limiting equations below we know that $\mathbf{V}^{-1}$ is bounded from above and strictly positive definite. The eigenvector matrix $\mathbf{E}$ of $\mathbf{V}^{-1}$ is that of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$ and hence, conditional on a given value of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$, not a function of t_c, t_s or σ_n^2/R.

[1] The latter found using the property that the maximum eigenvalue is less than the trace.

Next we substitute the notation in (I.1) into the equation for asymptotic simple beliefs to find

(I.4) $$\bar{\boldsymbol{\beta}}_s \xrightarrow{a.s.} \boldsymbol{\beta}_s + c\mathbf{M}^{-1}\boldsymbol{\beta}_s = \boldsymbol{\beta}_s + c\left[\mathbf{V} + \frac{\boldsymbol{\beta}_s\boldsymbol{\beta}_s'}{\boldsymbol{\beta}'\boldsymbol{\beta}}\right]^{-1}\boldsymbol{\beta}_s$$

$$= \boldsymbol{\beta}_s + c\left[\mathbf{V}^{-1} - \frac{\mathbf{V}^{-1}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'\mathbf{V}^{-1}/\boldsymbol{\beta}'\boldsymbol{\beta}}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}}\right]\boldsymbol{\beta}_s = \boldsymbol{\beta}_s + \frac{c\mathbf{V}^{-1}\boldsymbol{\beta}_s}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}},$$

Using this we see that when simple beliefs are proportional to $\boldsymbol{\beta}_s$ asymptotically only one of the a_i in (I.2) is non-zero, i.e., one of the eigenvectors in $\mathbf{E}$ is $\boldsymbol{\beta}_s/(\boldsymbol{\beta}_s'\boldsymbol{\beta}_s)^{½}$ and the rest are orthogonal to $\boldsymbol{\beta}_s$, as in this case:

(I.5) $$\alpha\boldsymbol{\beta}_s = \bar{\boldsymbol{\beta}}_s \xrightarrow{a.s.} \boldsymbol{\beta}_s + \frac{c\mathbf{V}^{-1}\boldsymbol{\beta}_s}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}} \Rightarrow \mathbf{V}^{-1}\boldsymbol{\beta}_s \xrightarrow{a.s.} \frac{(\alpha-1)(1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta})}{c}\boldsymbol{\beta}_s,$$

so $\boldsymbol{\beta}_s/(\boldsymbol{\beta}_s'\boldsymbol{\beta}_s)^{½}$ is an eigenvector of $\mathbf{V}^{-1}$. Alternatively, if $\lambda_i = \lambda$ for all $a_i \neq 0$, then since $\boldsymbol{a} = \mathbf{E}'\boldsymbol{\beta}_s$, we have $\mathbf{V}^{-1}\boldsymbol{\beta}_s = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}'\boldsymbol{\beta}_s = \mathbf{E}(\lambda\mathbf{I}_{k_s})\mathbf{a} = \lambda\mathbf{E}\mathbf{a} = \lambda\mathbf{E}\mathbf{E}'\boldsymbol{\beta}_s = \lambda\boldsymbol{\beta}_s$, so using (I.4) we again see that beliefs are proportional to $\boldsymbol{\beta}_s$

(I.6) $$\bar{\boldsymbol{\beta}}_s \xrightarrow{a.s.} \boldsymbol{\beta}_s + \frac{c\mathbf{V}^{-1}\boldsymbol{\beta}_s}{1+\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}} = \left[1 + \frac{c\lambda}{1+\lambda\boldsymbol{\beta}_s'\boldsymbol{\beta}_s/\boldsymbol{\beta}'\boldsymbol{\beta}}\right]\boldsymbol{\beta}_s.$$

In this case we can express the eigenbasis of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$ in such a way that only one a_i is non-zero. In sum, if beliefs are proportional to $\boldsymbol{\beta}_s$ only one a_i is non-zero, and if not then the eigenvalues cannot all be equal for all $a_i \neq 0$.

When the complex are in power t_c is the only element that changes in $\mathbf{V}$ and hence the asymptotic effect on (I.2a) and (I.2b) can be calculated by simply looking at the implied changes in the eigenvalues in (I.3), as the eigenvectors remain those of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$. When the simple are in power, t_s changes, with effects through eigenvalues similar to those of the complex, but $\mathbf{X}_{ss}'\mathbf{X}_{ss}$ also changes, with effects on both the eigenvalues and eigenvectors, i.e., the a_i terms in (I.2). We first calculate the effects of changes in t_c and t_s, and then examine the effects of changes in $\mathbf{X}_{ss}'\mathbf{X}_{ss}$, showing that they move (I.2a) and (I.2b) in the same direction as implied by increases in t_s.

Taking derivatives with respect to t_c and t_s, we have

(I.7) $$\frac{d\lambda_i}{dt_c} = \frac{R(\gamma_i + t_s\sigma_n^2)}{(\gamma_i + (t_s+t_c)\sigma_n^2)^2} > 0 \quad \text{and} \quad \frac{d\lambda_i}{dt_s} = -\frac{Rt_c\sigma_n^2}{(\gamma_i + (t_s+t_c)\sigma_n^2)^2} < 0.$$

From (I.7) we see that when the complex are in power t_c increases and all of the eigenvalues of $\mathbf{V}^{-1}$ increase (with no change in the eigenvectors), so $\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s$ increases and, consequently, so does $f(c\mathbf{M}^{-1})$. When the simple are in power t_s increases, which lowers all of the eigenvalues of $\mathbf{V}^{-1}$ (without changing the eigenvectors) and hence lowers $f(c\mathbf{M}^{-1})$. Taking the derivative of (I.2b) with respect to any eigenvalue, we find:

$$\text{(I.8)} \qquad \frac{\mathrm{d}\left(\dfrac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s}{(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2}\right)}{\mathrm{d}\lambda_i} = \frac{2a_i^2}{(\sum \lambda_i a_i^2)^3}\left[\lambda_i \sum_{j=1}^{k_s} \lambda_j a_j^2 - \sum_{j=1}^{k_s} \lambda_j^2 a_j^2\right]$$

So,

$$\text{(I.9)} \qquad \frac{\mathrm{d}\left(\dfrac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s}{(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2}\right)}{\mathrm{d}t_c} = \frac{2}{(\sum \lambda_i a_i^2)^3}\sum_{i=1}^{k_s} a_i^2\left[\lambda_i \sum_{j=1}^{k_s} \lambda_j a_j^2 - \sum_{j=1}^{k_s} \lambda_j^2 a_j^2\right]\frac{\mathrm{d}\lambda_i}{\mathrm{d}t_c}$$

$$= \frac{2}{(\sum \lambda_i a_i^2)^3}\sum_{i=1}^{k_s}\sum_{j=1}^{k_s} a_i^2 a_j^2(\lambda_i\lambda_j - \lambda_j^2)\frac{\mathrm{d}\lambda_i}{\mathrm{d}t_c}$$

$$= \frac{2}{(\sum \lambda_i a_i^2)^3}\sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s} a_i^2 a_j^2\left[(\lambda_i\lambda_j - \lambda_j^2)\frac{\mathrm{d}\lambda_i}{\mathrm{d}t_c} + (\lambda_i\lambda_j - \lambda_i^2)\frac{\mathrm{d}\lambda_j}{\mathrm{d}t_c}\right] \le 0,$$

as

$$\text{(I.10)} \qquad (\lambda_i\lambda_j - \lambda_j^2)\frac{\mathrm{d}\lambda_i}{\mathrm{d}t_c} + (\lambda_i\lambda_j - \lambda_i^2)\frac{\mathrm{d}\lambda_j}{\mathrm{d}t_c} = -\frac{\sigma_n^2(\gamma_i - \gamma_j)^2\lambda_i^3\lambda_j^3}{t_c^3 R^3} < 0,$$

with equality when $\sigma_n^2 = 0$ or a_i is non-zero for only one eigenvalue (i.e., the simple are on the level curve associated with the steady state with beliefs proportional to $\boldsymbol{\beta}_s$). Similarly,

$$\text{(I.11)} \qquad \frac{\mathrm{d}\left(\dfrac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s}{(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2}\right)}{\mathrm{d}t_s} = \frac{2}{(\sum \lambda_i a_i^2)^3}\sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s} a_i^2 a_j^2\left[(\lambda_i\lambda_j - \lambda_j^2)\frac{\mathrm{d}\lambda_i}{\mathrm{d}t_s} + (\lambda_i\lambda_j - \lambda_i^2)\frac{\mathrm{d}\lambda_j}{\mathrm{d}t_s}\right] \le 0$$

as

$$\text{(I.12)} \qquad (\lambda_i\lambda_j - \lambda_j^2)\frac{\mathrm{d}\lambda_i}{\mathrm{d}t_s} + (\lambda_i\lambda_j - \lambda_i^2)\frac{\mathrm{d}\lambda_j}{\mathrm{d}t_s} = -\frac{\sigma_n^2(\gamma_i - \gamma_j)^2\lambda_i^3\lambda_j^3}{t_c^3 R^3} < 0,$$

with, once again, equality when $\sigma_n^2 = 0$ or when beliefs are proportional to $\boldsymbol{\beta}_s$ and a_i is non-zero for only one eigenvalue. Intuition for why (I.9) and (I.11) are identical can be found by noting that while t_c appears in the numerator of (I.3), this element implicitly cancels in the ratio (I.2b). Consequently, all that is left is the influence of t_c and t_s in the denominator of (I.3), where they are both multiplied by σ_n^2. As time passes, regardless of which type is in power, random noise lowers the angle of the deviation of the simple's beliefs from the direction implied by the true parameter values.

We now consider the impact of periods when the simple are in power through its effects on $\mathbf{X}_{ss}'\mathbf{X}_{ss}$. $f(c\mathbf{M}^{-1})$ is monotonically increasing in $\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s$, with $\mathbf{V}$ as defined in (I.1). Each period when the simple are in power and implement policies $\mathbf{x}$ generates a rank one update of $\mathbf{V}$, so that $\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s$ becomes

$$\text{(I.13)} \qquad \boldsymbol{\beta}_s'\left[\mathbf{V}+\frac{\mathbf{xx}'}{t_cR}\right]^{-1}\boldsymbol{\beta}_s = \boldsymbol{\beta}_s'\left[\mathbf{V}^{-1}-\frac{\mathbf{V}^{-1}\mathbf{xx}'\mathbf{V}^{-1}/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR}\right]\boldsymbol{\beta}_s$$

$$= \boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s - \frac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{xx}'\mathbf{V}^{-1}\boldsymbol{\beta}_s/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR} < \boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s,$$

so this effect lowers $f(c\mathbf{M}^{-1})$ as does (as already proven) the increase in t_s that accompanies periods when the simple are in power.

Turning to the ratio $f(c^2\mathbf{M}^{-2})/f(c\mathbf{M}^{-1})^2$, equal to $\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s/(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2$ as shown in (I.1), we again calculate the effects of the rank-one update of $\mathbf{V}$

$$\text{(I.14)} \qquad \frac{\boldsymbol{\beta}_s'\left[\mathbf{V}^{-1}-\frac{\mathbf{V}^{-1}\mathbf{xx}'\mathbf{V}^{-1}/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR}\right]\left[\mathbf{V}^{-1}-\frac{\mathbf{V}^{-1}\mathbf{xx}'\mathbf{V}^{-1}/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR}\right]\boldsymbol{\beta}_s}{\boldsymbol{\beta}_s'\left[\mathbf{V}^{-1}-\frac{\mathbf{V}^{-1}\mathbf{xx}'\mathbf{V}^{-1}/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR}\right]\boldsymbol{\beta}_s\boldsymbol{\beta}_s'\left[\mathbf{V}^{-1}-\frac{\mathbf{V}^{-1}\mathbf{xx}'\mathbf{V}^{-1}/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR}\right]\boldsymbol{\beta}_s} =$$

$$\frac{m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR)^2 - 2(m^2_{\boldsymbol{\beta}_s'\mathbf{x}}m^1_{\boldsymbol{\beta}_s'\mathbf{x}}/t_cR)(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR) + m^1_{\boldsymbol{\beta}_s'\mathbf{x}}m^2_{\mathbf{x}'\mathbf{x}}m^1_{\boldsymbol{\beta}_s'\mathbf{x}}/(t_cR)^2}{[m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR) - m^1_{\boldsymbol{\beta}_s'\mathbf{x}}m^1_{\boldsymbol{\beta}_s'\mathbf{x}}/t_cR]^2}, \text{ with } m^i_{\mathbf{a}'\mathbf{b}} = \mathbf{a}'\mathbf{V}^{-i}\mathbf{b}.$$

We wish to show this is $\leq \boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{V}^{-1}\boldsymbol{\beta}_s/(\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s)^2 = m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}$, with equality only when $\bar{\boldsymbol{\beta}}_s$ is proportional to $\boldsymbol{\beta}_s$, i.e., when simple beliefs lie along the lowest level curve where $f(c^2\mathbf{M}^{-2}) = f(c\mathbf{M}^{-1})^2/\boldsymbol{\beta}'_s\boldsymbol{\beta}_s$. If $\bar{\boldsymbol{\beta}}_s$ is proportional to $\boldsymbol{\beta}_s$, then so is policy implemented by the simple. Say $\mathbf{x} = \alpha\boldsymbol{\beta}_s$, then we have $m^i_{\boldsymbol{\beta}_s'\mathbf{x}} = \alpha m^i_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}$ and (I.14) simplifies to:

$$\text{(I.15)} \qquad \frac{m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR)^2 - 2\alpha^2 m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/t_cR)(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR) + \alpha^4 m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/(t_cR)^2}{[m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR) - \alpha^2 m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/t_cR]^2}$$

$$= \frac{m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}[(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR) - \alpha^2 m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/t_cR]^2}{m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}[(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR) - \alpha^2 m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/t_cR]^2} = \frac{m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}}{m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}},$$

as desired. Our next task is to show that (I.14) is asymptotically strictly less than $m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}$ if beliefs are not proportional to $\boldsymbol{\beta}_s$.

We begin by using $\mathbf{x} = \bar{\boldsymbol{\beta}}_s\sqrt{R/\bar{\boldsymbol{\beta}}_s'\bar{\boldsymbol{\beta}}_s}$ and (I.4) to find that

(I.16) $$m^i_{\boldsymbol{\beta}'_s\mathbf{x}} = \sqrt{\frac{R}{\bar{\boldsymbol{\beta}}'_s\bar{\boldsymbol{\beta}}_s}}\left[m^i_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} + \frac{c\,m^{i+1}_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})}\right]$$

$$\text{and } m^i_{\mathbf{x}'\mathbf{x}} = \frac{R}{\bar{\boldsymbol{\beta}}'_s\bar{\boldsymbol{\beta}}_s}\left[m^i_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} + \frac{2c\,m^{i+1}_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})} + \frac{c^2 m^{i+2}_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})^2}\right].$$

(I.16) tells us that all $m^i_{\boldsymbol{\beta}'_s\mathbf{x}}$ and $m^i_{\mathbf{x}'\mathbf{x}}$ can be expressed as a combination of $m^j_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}$ terms. Each $m^j_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}$ is asymptotically bounded, as

(I.17) $$\boldsymbol{\beta}'_s\mathbf{V}^{-j}\boldsymbol{\beta}_s \le \lambda_{\max}(\mathbf{V}^{-j})\boldsymbol{\beta}'_s\boldsymbol{\beta}_s = \frac{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}{\lambda_{\min}(\mathbf{V})^j} \le \frac{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}{\lambda_{\min}\left(\mathbf{V}-\mathbf{X}'_{ss}\mathbf{X}_{ss}/t_cR\right)^j} \le \frac{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}{(\sigma_n^2/R)^j}$$

where we have made use of the definition of $\mathbf{V}$ from (I.1). Added to that the fact that (I.4) implies that $\bar{\boldsymbol{\beta}}'_s\bar{\boldsymbol{\beta}}_s \ge \boldsymbol{\beta}'_s\boldsymbol{\beta}_s$, and we can see that all $m^i_{\boldsymbol{\beta}'_s\mathbf{x}}$ and $m^i_{\mathbf{x}'\mathbf{x}}$ are bounded from above and the limit of (I.14) as t_c goes to infinity is $m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}$, as should be expected since the rank one updates of $\mathbf{V}$, $\mathbf{x}/(t_cR)^{1/2}$, get smaller and smaller.

With the preceding in mind, consider (I.14) as a function of t_c, $g(t_c)$, with[2]

(I.18) $$g'(t_c) = \left(\boldsymbol{\beta}'_s\left[\mathbf{V}^{-1} - \frac{\mathbf{V}^{-1}\mathbf{x}\mathbf{x}'\mathbf{V}^{-1}/t_cR}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_cR}\right]\boldsymbol{\beta}_s\right)^{-3}\frac{2m^1_{\boldsymbol{\beta}'_s\mathbf{x}}}{t_c^2R(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR)^3}$$

$$*\left(\underbrace{\frac{m^1_{\boldsymbol{\beta}'_s\mathbf{x}}[m^2_{\boldsymbol{\beta}'_s\mathbf{x}}m^1_{\boldsymbol{\beta}'_s\mathbf{x}} - m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^2_{\mathbf{x}'\mathbf{x}}]}{t_cR}}_{c_1} + \underbrace{(1+m^1_{\mathbf{x}'\mathbf{x}}/t_cR)}_{c_2}\underbrace{[m^2_{\boldsymbol{\beta}'_s\mathbf{x}}m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}'_s\mathbf{x}}]}_{c_3}\right).$$

Substituting using (I.16), we have

(I.19) $$c_3 = \sqrt{\frac{R}{\bar{\boldsymbol{\beta}}'_s\bar{\boldsymbol{\beta}}_s}}\left(\left[m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} + \frac{c\,m^3_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})}\right]m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}\left[m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} + \frac{c\,m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})}\right]\right)$$

$$= \sqrt{\frac{R}{\bar{\boldsymbol{\beta}}'_s\bar{\boldsymbol{\beta}}_s}}\frac{c(m^3_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s})}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})} \ge 0,$$

$$\text{as } m^3_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} = \boldsymbol{\beta}'_s\mathbf{V}^{-3}\boldsymbol{\beta}_s\boldsymbol{\beta}'_s\mathbf{V}^{-1}\boldsymbol{\beta}_s = (\mathbf{V}^{-3/2}\boldsymbol{\beta}_s)'(\mathbf{V}^{-3/2}\boldsymbol{\beta}_s)(\mathbf{V}^{-1/2}\boldsymbol{\beta}_s)'(\mathbf{V}^{-1/2}\boldsymbol{\beta}_s)$$

$$\ge (\mathbf{V}^{-3/2}\boldsymbol{\beta}_s)'(\mathbf{V}^{-1/2}\boldsymbol{\beta}_s)(\mathbf{V}^{-3/2}\boldsymbol{\beta}_s)'(\mathbf{V}^{-1/2}\boldsymbol{\beta}_s) = \boldsymbol{\beta}'_s\mathbf{V}^{-2}\boldsymbol{\beta}_s\boldsymbol{\beta}'_s\mathbf{V}^{-2}\boldsymbol{\beta}_s = m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s},$$

$$\text{while } c_1 = \frac{m^1_{\boldsymbol{\beta}'_s\mathbf{x}}}{t_cR}\frac{R}{\bar{\boldsymbol{\beta}}'_s\bar{\boldsymbol{\beta}}_s}\left[\frac{c(m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} - m^3_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s})}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})} + \frac{c^2(m^3_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s} - m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}m^4_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s})}{(1+m^1_{\boldsymbol{\beta}'_s\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})^2}\right],$$

where we once again use the Cauchy-Schwarz inequality. We are unable to sign c_1, but since c_2

[2]We ignore the effect of t_c on $\mathbf{V}$ as we are trying to prove the sign of the rank one update $\mathbf{x}$ given $\mathbf{V}$.

> 1 and $c_3 \geq 0$, if c_1 is strictly positive it follows that $g'(t_c)$ is strictly positive and consequently $g(t_c)$ is strictly less than $m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} / m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}$ for finite t_c as long as simple beliefs are not proportional to $\boldsymbol{\beta}_s$. Going forward, we assume this is not the case, i.e., that $c_1 \leq 0$.

Using the work above, we formally note the upper bounds on $R / \overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s$, $m^1_{\boldsymbol{\beta}_s'\mathbf{x}}$ and the maximum eigenvalue of $\mathbf{V}^{-1}$

(I.20) $$\frac{R}{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} \geq \frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s},\ m^{1*}_{\boldsymbol{\beta}_s'\mathbf{x}} = \sqrt{R\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}\left[\frac{R}{\sigma_n^2} + c\left(\frac{R}{\sigma_n^2}\right)^2\right] \geq m^1_{\boldsymbol{\beta}_s'\mathbf{x}},\ \&\ \lambda^* = \frac{R}{\sigma_n^2} \geq \lambda_{\max}[\mathbf{V}^{-1}],$$

and define t^* as

(I.21) $$t^* = 2\sqrt{\frac{R}{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}}\frac{m^{1*}_{\boldsymbol{\beta}_s'\mathbf{x}}}{R}\max(1, c\lambda^*).$$

Substituting into $c_1 + c_2c_3$ using (I.20) and $t_c > t^*$

(I.22) $$\underbrace{\frac{m^1_{\boldsymbol{\beta}_s'\mathbf{x}}[m^2_{\boldsymbol{\beta}_s'\mathbf{x}}m^1_{\boldsymbol{\beta}_s'\mathbf{x}} - m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\mathbf{x}'\mathbf{x}}]}{t_cR}}_{\leq \text{ by assumption}} + \underbrace{(1 + m^1_{\mathbf{x}'\mathbf{x}}/t_cR)}_{>1}\underbrace{[m^2_{\boldsymbol{\beta}_s'\mathbf{x}}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\mathbf{x}}]}_{\geq 0 \text{ from (I.19)}}$$

$$\geq \frac{m^1_{\boldsymbol{\beta}_s'\mathbf{x}}}{t^*R}\frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}\left[\frac{c(m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s})}{(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})} + \frac{c^2(m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^4_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s})}{(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})^2}\right]$$

$$+ \sqrt{\frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}}\frac{c(m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s})}{(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})}$$

$$= \underbrace{\sqrt{\frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}}\frac{c(m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s})}{(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})}}_{\geq 0 \text{ from (I.19)}}\underbrace{\left[\frac{1}{2} - \sqrt{\frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}}\frac{m^1_{\boldsymbol{\beta}_s'\mathbf{x}}}{t^*R}\right]}_{\geq 0 \text{ from (I.21)}} +$$

$$\frac{m^1_{\boldsymbol{\beta}_s'\mathbf{x}}c^2}{t^*R(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})^2}\frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}\left[(m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^4_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}) + \left(\begin{array}{c}\underbrace{\sqrt{\frac{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}{R}}\frac{t^*R(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})}{2cm^1_{\boldsymbol{\beta}_s'\mathbf{x}}}}_{\geq \lambda^* \text{ from (I.21)}} \\ * \underbrace{(m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s})}_{\geq 0 \text{ from (I.19)}}\end{array}\right)\right]$$

$$\geq \frac{m^1_{\boldsymbol{\beta}_s'\mathbf{x}}c^2}{t^*R(1 + m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}/\boldsymbol{\beta}'\boldsymbol{\beta})^2}\frac{R}{\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s}\underbrace{\left[m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^4_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} + \lambda^*(m^3_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} - m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s})\right]}_{c_4}$$

Focusing on c_4 in the last line, as $m^i_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} = \boldsymbol{\beta}_s'\mathbf{V}^{-i}\boldsymbol{\beta}_s$, we use the spectral decomposition of $\mathbf{V}^{-1}$, as in (I.2) earlier

(I.23) $$c_4 = \sum_{i=1}^{k_s}\lambda_i^3 a_i^2 \sum_{i=1}^{k_s}\lambda_i^2 a_i^2 - \sum_{i=1}^{k_s}\lambda_i a_i^2 \sum_{i=1}^{k_s}\lambda_i^4 a_i^2 + \lambda^*\left[\sum_{i=1}^{k_s}\lambda_i^3 a_i^2 \sum_{i=1}^{k_s}\lambda_i a_i^2 - \sum_{i=1}^{k_s}\lambda_i^2 a_i^2 \sum_{i=1}^{k_s}\lambda_i^2 a_i^2\right]$$

$$= 2\sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s}(\lambda_i^3\lambda_j^2 - \lambda_i\lambda_j^4)a_i^2 a_j^2 + 2\lambda^*\sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s}(\lambda_i^3\lambda_j - \lambda_i^2\lambda_j^2)a_i^2 a_j^2$$

$$= 2\sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s}[\lambda_i^3\lambda_j^2 - \lambda_i\lambda_j^4 + \lambda^*(\lambda_i^3\lambda_j - \lambda_i^2\lambda_j^2)]a_i^2 a_j^2$$

$$\geq \sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s}[\lambda_i^3\lambda_j^2 - \lambda_i\lambda_j^4 + \lambda_i(\lambda_i^3\lambda_j - \lambda_i^2\lambda_j^2)]a_i^2 a_j^2 = \sum_{i=1}^{k_s-1}\sum_{j=i+1}^{k_s}[\lambda_i^4\lambda_j - \lambda_i\lambda_j^4]a_i^2 a_j^2 \ \geq\ 0,$$

where we have used the fact that the λ_i are ordered in decreasing order, with $\lambda_1 \geq ... \geq \lambda_i ... \geq \lambda_{k_s}$. The last line of (I.23) holds with strict inequality whenever there exists a difference between the maximum and minimum eigenvalues corresponding to non-zero a_i, i.e., simple beliefs are not proportional to $\boldsymbol{\beta}_s$. Consequently, we may conclude that for all $t_c > t^*$, as long as simple beliefs are not proportional to $\boldsymbol{\beta}_s$, $g'(t_c)$ is strictly positive and hence $g(t_c)$ is strictly less than $m^2_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} / m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s} m^1_{\boldsymbol{\beta}_s'\boldsymbol{\beta}_s}$. This concludes our proof that the rank one update of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$ when the simple are in power lowers the ratio $f(c^2\mathbf{M}^{-2})/f(c\mathbf{M}^{-1})^2$ as long as simple beliefs are not proportional to $\boldsymbol{\beta}_s$, i.e., as long as the economy is not on the (lowest) level curve in Figure A2 associated with the steady state.

To summarize, when the complex are in power, t_c increases in the formula for $\mathbf{M}$, which increases $f(c\mathbf{M}^{-1})$ and lowers the ratio $f(c^2\mathbf{M}^{-2})/f(c\mathbf{M}^{-1})^2$. When the simple are in power, t_s increases and there is also a rank-one update of $\mathbf{M}$ based upon implemented simple policy. Both of these lower both $f(c\mathbf{M}^{-1})$ and $f(c^2\mathbf{M}^{-2})/f(c\mathbf{M}^{-1})^2$. These are the results stated in (A28a) and (A28b). Turning to (A28c), we begin by noting that since the sum of the eigenvalues of a matrix equals the trace, the individual eigenvalues γ_i of $\mathbf{X}_{ss}'\mathbf{X}_{ss}$ are bounded from above by Rt_s. Consequently, we can bound the derivatives in (I.7) and prove that their limit is zero

(I.24) $$0 < \frac{d\lambda_i}{dt_c} = \frac{R(\gamma_i + t_s\sigma_n^2)}{(\gamma_i + (t_s+t_c)\sigma_n^2)^2} < \frac{R(Rt_s + t_s\sigma_n^2)}{t^2\sigma_n^4} < \frac{R(R+\sigma_n^2)}{t\sigma_n^4}$$

$$\text{and}\quad 0 > \frac{d\lambda_i}{dt_s} = -\frac{Rt_c\sigma_n^2}{(\gamma_i + (t_s+t_c)\sigma_n^2)^2} > -\frac{R}{t\sigma_n^2}$$

$$\Rightarrow\quad 0 \leq \lim_{t\to\infty}\frac{d\lambda_i}{dt_c} \leq \lim_{t\to\infty}\frac{R(R+\sigma_n^2)}{t\sigma_n^4} = 0 \ \ \&\ \ 0 \geq \lim_{t\to\infty}\frac{d\lambda_i}{dt_s} \geq \lim_{t\to\infty} -\frac{R}{t\sigma_n^2} = 0.$$

The only remaining effect on $f(c\mathbf{M}^{-1})$ with the passage of time is through the rank one update of $\boldsymbol{\beta}_s'\mathbf{V}^{-1}\boldsymbol{\beta}_s$, which, as described earlier in (I.13), generates a change

(I.25)
$$-\frac{\boldsymbol{\beta}_s'\mathbf{V}^{-1}\mathbf{x}\mathbf{x}'\mathbf{V}^{-1}\boldsymbol{\beta}_s / t_c R}{1+\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}/t_c R} = -\frac{m^1_{\boldsymbol{\beta}_s'\mathbf{x}} m^1_{\boldsymbol{\beta}_s'\mathbf{x}} / t_c R}{1+m^1_{\mathbf{x}'\mathbf{x}} / t_c R}.$$

However, as shown in (I.16) and (I.17), all $m^i_{\boldsymbol{\beta}_s'\mathbf{x}}$ and $m^i_{\mathbf{x}'\mathbf{x}}$ are bounded from above, while we established much earlier above that t_c goes to infinity (outside of equilibrium paths of probability measure zero which we are not examining). Consequently, the change in $f(c\mathbf{M}^{-1})$ through this mechanism goes to zero as well. This proves (A28c) and completes the proof of the convergence of $\overline{\boldsymbol{\beta}}_i$ and $\theta_i = t_i/t$ in this appendix.

II: Random Outcomes and the Political Cycle

As noted in the paper, a characteristic of political life seems to be that random outcomes benefit or harm incumbent parties. In this appendix we show that this feature arises in our model through the fully rational Bayesian updating of beliefs. Random shocks change estimates of the effectiveness of policy, but these effects are stronger for the incumbent party which is implementing its desired policy combination. Consequently, although the long run equilibrium involves cycles with each type on average in power for a determinate share of the time, a random negative shock to y lowers the voting intensity of incumbent groups relative to their opposition, hastening regime change, while random positive shocks to y strengthen the political position of incumbents, lengthening their stay in power in the current political cycle. The proof below shows that these statements, which form Proposition 1 in the text, are true in the probability limit.

To allow an examination of period by period beliefs, we introduce notation with respect to time, with the t x k matrix $\mathbf{H}_t$ denoting the history of policy up to time t, the vector $\mathbf{h}'_t$ the t^{th} row thereof, and $\mathbf{H}_{it}$ and $\mathbf{h}'_{it}$ the corresponding histories and t^{th} period policies that type i deems relevant. We focus on outcomes in the vicinity of the steady state and, to simplify the analysis, with negligible amounts of policy noise. The analysis below is in the context of the generalized model described in the appendix of the paper and makes use of the Lemmas and methods of proof therein.

The formula for mean Bayesian beliefs, based as it is upon regression coefficients, allows a simple representation of the updating of beliefs from period t to $t+1$

$$
\begin{aligned}
\text{(II.1)}\qquad \overline{\boldsymbol{\beta}}_{it+1} &= (\mathbf{H}'_{it+1}\mathbf{H}_{it+1})^{-1}\mathbf{H}'_{it+1}\mathbf{y}_{t+1} \\
&= \left[(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1} - \frac{(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}\mathbf{h}'_{it+1}(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}}{1+\mathbf{h}'_{it+1}(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}}\right](\mathbf{H}'_{it}\mathbf{y}_t + \mathbf{h}_{it+1}y_{t+1}) \\
&= \overline{\boldsymbol{\beta}}_{it} - \frac{(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}\mathbf{h}'_{it+1}\overline{\boldsymbol{\beta}}_{it}}{1+\mathbf{h}'_{it+1}(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}} + \frac{(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}y_{t+1}}{1+\mathbf{h}'_{it+1}(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}} \\
&= \overline{\boldsymbol{\beta}}_{it} + \frac{(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}}{1+\mathbf{h}'_{it+1}(\mathbf{H}'_{it}\mathbf{H}_{it})^{-1}\mathbf{h}_{it+1}}[y_{t+1} - \mathbf{h}'_{it+1}\overline{\boldsymbol{\beta}}_{it}],
\end{aligned}
$$

where in the second line we make use of the Sherman-Morrison formula for the rank one update of a matrix inverse. The term in brackets [] in the last line is the period $t+1$ prediction error based upon beliefs at the end of period t.

Asymptotically, the change in intensity of each type, $\overline{\boldsymbol\beta}'_{it+1}\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}'_{it}\overline{\boldsymbol\beta}_{it}$ = $(\overline{\boldsymbol\beta}_{it+1} + \overline{\boldsymbol\beta}_{it})'(\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}_{it})$, almost surely converges to $2\overline{\boldsymbol\beta}'_{it}(\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}_{it})$, since, using (II.1)

(II.2) $$(\overline{\boldsymbol\beta}_{it+1} + \overline{\boldsymbol\beta}_{it})'(\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}_{it}) - 2\overline{\boldsymbol\beta}'_{it}(\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}_{it}) = (\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}_{it})'(\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}_{it})$$

$$= \frac{\dfrac{\mathbf{h}'_{it+1}}{\sqrt{t}}\left(\dfrac{\mathbf{H}'_{it}\mathbf{H}_{it}}{t}\right)^{-2}\dfrac{\mathbf{h}_{it+1}}{\sqrt{t}}}{\left[1+\dfrac{\mathbf{h}'_{it+1}}{\sqrt{t}}\left(\dfrac{\mathbf{H}'_{it}\mathbf{H}_{it}}{t}\right)^{-1}\dfrac{\mathbf{h}_{it+1}}{\sqrt{t}}\right]^2}\left[\frac{y_{t+1} - \mathbf{h}'_{it+1}\overline{\boldsymbol\beta}_{it}}{\sqrt{t}}\right]^2 \xrightarrow{a.s.} 0.$$

The almost sure limit on the last line follows from the fact that for j = 1 or 2

(II.3) (a) $$\frac{\mathbf{h}'_{it+1}}{\sqrt{t}}\left(\frac{\mathbf{H}'_{it}\mathbf{H}_{it}}{t}\right)^{-j}\frac{\mathbf{h}_{it+1}}{\sqrt{t}} \le \frac{\dfrac{\mathbf{h}'_{it+1}\mathbf{h}_{it+1}}{t}}{\lambda_{\min}\left(\dfrac{\mathbf{H}'_{it}\mathbf{H}_{it}}{t}\right)^j} \le \frac{\dfrac{R}{t}+\dfrac{\mathbf{x}'_{it}\mathbf{n}_{it}}{t}+\dfrac{\mathbf{n}'_{it}\mathbf{x}_{it}}{t}+\dfrac{\mathbf{n}'_{it}}{\sqrt{t}}\dfrac{\mathbf{n}_{it}}{\sqrt{t}}}{\lambda_{\min}\left(\dfrac{\mathbf{H}'_{it}\mathbf{H}_{it}-\mathbf{X}'_{it}\mathbf{X}_{it}}{t}\right)^j} \xrightarrow{a.s.} \frac{0}{\sigma_n^{2j}},$$

(b) $$\frac{y_{t+1} - \mathbf{h}'_{it+1}\overline{\boldsymbol\beta}_{it}}{\sqrt{t}} = \frac{\mathbf{x}'_{t+1}\boldsymbol\beta - \mathbf{x}'_{it+1}\overline{\boldsymbol\beta}_{it}}{\sqrt{t}} + \frac{\mathbf{n}'_{t+1}\boldsymbol\beta - \mathbf{n}'_{it+1}\overline{\boldsymbol\beta}_{it}}{\sqrt{t}} + \frac{\varepsilon_{t+1}}{\sqrt{t}} \xrightarrow{a.s.} 0.$$

For the denominator of (II.3a) we use Lemma (A1) of the appendix to see that

(II.4) $$\frac{\mathbf{H}'_{it}\mathbf{H}_{it} - \mathbf{X}'_{it}\mathbf{X}_{it}}{t} = \frac{\mathbf{X}'_{it}\mathbf{N}_{it} + \mathbf{N}'_{it}\mathbf{X}_{it} + \mathbf{N}'_{it}\mathbf{N}_{it}}{t} \xrightarrow{a.s.} \sigma_n^2 I_{k_i}.$$

For the remaining terms we use the fact that policies $\mathbf{x}$ and beliefs $\overline{\boldsymbol\beta}_i$ converge almost surely to finite constants while, with v_t denoting either ε_{t+1} or any element of $\mathbf{n}_{it}$ or $\mathbf{n}_{it+1}$ divided by $\sqrt{t}$, we have $E(v_t^4) = \mu_4 / t^2$, where, following the assumption given in the paper's appendix, μ_4 is the bounded fourth moment of ε_{t+1} or any of the iid elements of $\mathbf{n}_{it}$. Consequently, applying Markov's Inequality we have for any $a > 0$

(II.5) $$P(v_t^4 \ge a^4) \le \frac{E(v_t^4)}{a^4} \Rightarrow P(|v_t| > a) \le \frac{\mu_4}{a^4 t^2},$$

so, as discussed in the paper's appendix, from the Borel-Cantelli Lemma we know that v_t almost surely converges to 0.

Using (II.2) and (II.1), we can say that asymptotically the change in intensity of type i is given by:

(II.6) $$\overline{\boldsymbol\beta}'_{it+1}\overline{\boldsymbol\beta}_{it+1} - \overline{\boldsymbol\beta}'_{it}\overline{\boldsymbol\beta}_{it} \xrightarrow{a.s.} \frac{2\overline{\boldsymbol\beta}'_{it}\left(\dfrac{\mathbf{H}'_{it}\mathbf{H}_{it}}{t}\right)^{-1}\dfrac{\mathbf{h}_{it+1}}{\sqrt{t}}}{1+\dfrac{\mathbf{h}'_{it+1}}{\sqrt{t}}\left(\dfrac{\mathbf{H}'_{it}\mathbf{H}_{it}}{t}\right)^{-1}\dfrac{\mathbf{h}_{it+1}}{\sqrt{t}}}\frac{[y_{t+1} - \mathbf{h}'_{it+1}\overline{\boldsymbol\beta}_{it}]}{\sqrt{t}}.$$

Since beliefs almost surely converge, it is of course easily shown that the term on the right-hand side converges almost surely to 0. Instead of examining this degenerate case, we shall consider the probability limit of t times the change in intensity. The shift in emphasis from almost sure to probability limit comes from the fact that if for random variables a, b and c we have $a \xrightarrow{a.s.} bc$ and $b \xrightarrow{a.s.} b^*$, it is not necessarily true that $a \xrightarrow{a.s.} b^*c$ if c is unbounded, but it is true that $a \xrightarrow{p} b^*c$ if the second moment of c is bounded. To see this, note that if $a \xrightarrow{a.s.} bc$ and $b \xrightarrow{a.s.} b^*$ then outside of paths of probability measure zero for every $\varepsilon > 0$ and $\delta > 0$ there exists a $t_{\varepsilon,\delta}$ sufficiently large such that $|a - bc| < \varepsilon/2$ and $|b - b^*| < (\varepsilon/2)\sqrt{\delta/E(c^2)}$ for all $t > t_{\varepsilon,\delta}$. We can then say that for all $t > t_{\varepsilon,\delta}$ along such paths

(II.7)
$$P(|a - b^*c| > \varepsilon) \le P(|a - bc| + |b - b^*||c| > \varepsilon)$$
$$\le P(|c| > \sqrt{E(c^2)/\delta}) = P(c^2 > E(c^2)/\delta) \le \delta.$$

where in the last we use Markov's Inequality. Since the remaining paths are of probability measure zero, this establishes that $a \xrightarrow{p} b^*c$.[3]

We now consider the limits of various elements on the right-hand side of (II.6). As shown in (II.3a), the quadratic form in the denominator of (II.6) almost surely converges to zero. As for the term $(\mathbf{H}'_{it}\mathbf{H}_{it}/t)^{-1}$ in the numerator, we use the fact that policies and the share of time each type spends in power almost surely converge to steady state values, and the almost sure limits of Lemma (A1), to say that

(II.8)
$$\frac{\mathbf{H}'_{ct}\mathbf{H}_{ct}}{t} \xrightarrow{a.s.} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{bmatrix} \text{ and } \frac{\mathbf{H}'_{st}\mathbf{H}_{st}}{t} \xrightarrow{a.s.} \begin{bmatrix} \mathbf{A} & \mathbf{0}_{k_b x k_{\sim c}} \\ \mathbf{0}_{k_{\sim c} x k_b} & \sigma_n^2 \mathbf{I}_{k_{\sim c}} \end{bmatrix},$$

$$\text{with } \mathbf{A} = (\theta_s \tau^{*2} + \theta_c)\boldsymbol{\beta}_b\boldsymbol{\beta}'_b \frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}} + \sigma_n^2\mathbf{I}_{k_b},\ \mathbf{B} = \theta_c\boldsymbol{\beta}_b\boldsymbol{\beta}'_{\sim s}\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}},\ \&\ \mathbf{C} = \theta_c\boldsymbol{\beta}_{\sim s}\boldsymbol{\beta}'_{\sim s}\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}} + \sigma_n^2\mathbf{I}_{k_{\sim s}},$$

where we use the subscript $\sim i$ to denote policies that each type deems irrelevant, subscript b to denote the policies both deem relevant, $k_{\sim i}$ and k_b the number of such policies, and make use of the fact that simple beliefs and policies in areas the complex deem irrelevant ($\sim c$) converge to the true parameter values of 0. With regards to the prediction error, since beliefs and policies

[3]To understand the distinction between this result and the almost sure limit, note that if c is an unbounded iid random variable and b does not converge to b^* quickly enough, then for every $\varepsilon > 0$ on a positive measure of paths there occur an infinite, albeit increasingly rare, number of events were a deviates from b^*c by more than ε, and hence a does not converge almost surely to that limit. If it can be established that b converges sufficiently rapidly to b^*, then with an appeal to the Borel-Cantelli Lemma and bounds on the higher moments of c it is possible to establish almost sure convergence, but the probability limit is sufficient for our purposes.

converge almost surely, and the second moments of the noise and output shocks are bounded, we can say that:

$$\text{(II.9)} \qquad y_{t+1} - \mathbf{h}'_{st+1}\overline{\boldsymbol{\beta}}_{st} \xrightarrow{p} \mathbf{h}'_{t+1}\boldsymbol{\beta} - \mathbf{h}'_{st+1}\tau^{*}\boldsymbol{\beta}_{s} + \varepsilon_{t+1} \;\;\&\;\; y_{t+1} - \mathbf{h}'_{ct+1}\overline{\boldsymbol{\beta}}_{ct} \xrightarrow{p} \varepsilon_{t+1}.$$

These results allow us to express the asymptotic change in intensity as

$$\text{(II.10)} \qquad t(\overline{\boldsymbol{\beta}}'_{st+1}\overline{\boldsymbol{\beta}}_{st+1} - \overline{\boldsymbol{\beta}}'_{st}\overline{\boldsymbol{\beta}}_{st}) \xrightarrow{p} 2\tau^{*}\boldsymbol{\beta}'_{s}\begin{bmatrix} \mathbf{A} & \mathbf{0}_{k_b x k_{\sim c}} \\ \mathbf{0}_{k_{\sim c} x k_b} & \sigma_n^2\mathbf{I}_{k_{\sim c}} \end{bmatrix}^{-1}\mathbf{h}_{st+1}[\mathbf{h}'_{t+1}\boldsymbol{\beta} - \mathbf{h}'_{st+1}\tau^{*}\boldsymbol{\beta}_{s} + \varepsilon_{t+1}],$$

$$t(\overline{\boldsymbol{\beta}}'_{ct+1}\overline{\boldsymbol{\beta}}_{ct+1} - \overline{\boldsymbol{\beta}}'_{ct}\overline{\boldsymbol{\beta}}_{ct}) \xrightarrow{p} 2\boldsymbol{\beta}'_{c}\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{bmatrix}^{-1}\mathbf{h}_{ct+1}\varepsilon_{t+1}.$$

Finally, we note the formula for a block matrix inverse and calculate the limits of some useful quadratic forms as policy noise goes to zero:

$$\text{(II.11)} \qquad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1} & -(\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{C}^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}'(\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1} & \mathbf{C}^{-1}+\mathbf{C}^{-1}\mathbf{B}'(\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{C}^{-1} \end{bmatrix},$$

$$\lim_{\sigma_n^2\to 0}\boldsymbol{\beta}'_b\mathbf{A}^{-1}\boldsymbol{\beta}_b = \lim_{\sigma_n^2\to 0}\boldsymbol{\beta}'_b\left[\frac{1}{\sigma_n^2}\mathbf{I}_{k_b} - \frac{(\theta_s\tau^{*2}+\theta_c)\boldsymbol{\beta}_b\boldsymbol{\beta}'_b(R/\boldsymbol{\beta}'\boldsymbol{\beta})/\sigma_n^4}{1+(\theta_s\tau^{*2}+\theta_c)\boldsymbol{\beta}'_b\boldsymbol{\beta}_b(R/\boldsymbol{\beta}'\boldsymbol{\beta})/\sigma_n^2}\right]\boldsymbol{\beta}_b$$

$$= \lim_{\sigma_n^2\to 0}\frac{\boldsymbol{\beta}'_b\boldsymbol{\beta}_b}{\sigma_n^2+(\theta_s\tau^{*2}+\theta_c)\boldsymbol{\beta}'_b\boldsymbol{\beta}_b(R/\boldsymbol{\beta}'\boldsymbol{\beta})} = \frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{(\theta_s\tau^{*2}+\theta_c)R},$$

$$\lim_{\sigma_n^2\to 0}\boldsymbol{\beta}'_{\sim s}\mathbf{C}^{-1}\boldsymbol{\beta}_{\sim s} = \lim_{\sigma_n^2\to 0}\boldsymbol{\beta}'_{\sim s}\left[\frac{1}{\sigma_n^2}\mathbf{I}_{k_{\sim s}} - \frac{\theta_c\boldsymbol{\beta}_{\sim s}\boldsymbol{\beta}'_{\sim s}(R/\boldsymbol{\beta}'\boldsymbol{\beta})/\sigma_n^4}{1+\theta_c\boldsymbol{\beta}'_{\sim s}\boldsymbol{\beta}_{\sim s}(R/\boldsymbol{\beta}'\boldsymbol{\beta})/\sigma_n^2}\right]\boldsymbol{\beta}_{\sim s}$$

$$= \lim_{\sigma_n^2\to 0}\frac{\boldsymbol{\beta}'_{\sim s}\boldsymbol{\beta}_{\sim s}}{\sigma_n^2+\theta_c\boldsymbol{\beta}'_{\sim s}\boldsymbol{\beta}_{\sim s}(R/\boldsymbol{\beta}'\boldsymbol{\beta})} = \frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{\theta_c R},$$

$$\lim_{\sigma_n^2\to 0}\boldsymbol{\beta}'_{\sim c}\left[\sigma_n^2\mathbf{I}_{k_{\sim c}}\right]^{-1}\boldsymbol{\beta}_{\sim c} = \lim_{\sigma_n^2\to 0}\boldsymbol{\beta}'_{\sim c}\boldsymbol{\beta}_{\sim c}/\sigma_n^2 = \lim_{\sigma_n^2\to 0} 0/\sigma_n^2 = 0.$$

Since we are considering the limit as the variance of policy noise goes to zero, we also take $\mathbf{h}_{t+1}$ in (II.10) as equal to $\mathbf{x}_{t+1}$, the intentional policy vector of that time period.

Asymptotically the simple implement policies $\tau^{*}\boldsymbol{\beta}_s\sqrt{R/\boldsymbol{\beta}'\boldsymbol{\beta}}$ for the policies they believe are relevant and $\mathbf{0}_{k_{\sim s}x1}$ for those they believe are irrelevant, so using the preceding results the change in the intensity of both types when the simple are in power is

$$\text{(II.12)}\qquad \lim_{\sigma_n^2 \to 0} \operatorname{plim}\ t(\overline{\boldsymbol{\beta}}'_{st+1}\overline{\boldsymbol{\beta}}_{st+1} - \overline{\boldsymbol{\beta}}'_{st}\overline{\boldsymbol{\beta}}_{st}) = \lim_{\sigma_n^2 \to 0} 2\tau^*\left[\boldsymbol{\beta}'_b\mathbf{A}^{-1}\boldsymbol{\beta}_b + \boldsymbol{\beta}'_{\sim c}\boldsymbol{\beta}_{\sim c}/\sigma_n^2\right]\tau^*\sqrt{\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}}}$$

$$*\left[\tau^*\boldsymbol{\beta}'_s\sqrt{\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}}}\boldsymbol{\beta}_s - \tau^*\boldsymbol{\beta}'_s\sqrt{\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}}}\tau^*\boldsymbol{\beta}_s + \varepsilon_{t+1}\right] = \frac{2\boldsymbol{\beta}'\boldsymbol{\beta}\tau^*[1-\tau^*]}{\theta_s\tau^{*2}+\theta_c} + \sqrt{\frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{R}}\frac{2\tau^{*2}\varepsilon_{t+1}}{\theta_s\tau^{*2}+\theta_c},$$

$$\lim_{\sigma_n^2 \to 0} \operatorname{plim} t(\overline{\boldsymbol{\beta}}'_{ct+1}\overline{\boldsymbol{\beta}}_{ct+1} - \overline{\boldsymbol{\beta}}'_{ct}\overline{\boldsymbol{\beta}}_{ct}) = \lim_{\sigma_n^2 \to 0} 2\left[\boldsymbol{\beta}'_b \quad \boldsymbol{\beta}'_{\sim s}\right]\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}' & \mathbf{C}\end{bmatrix}^{-1}\begin{bmatrix}\tau^*\boldsymbol{\beta}_b\\ \mathbf{0}_{k_{\sim s}x1}\end{bmatrix}\sqrt{\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}}}\varepsilon_{t+1}$$

$$= \lim_{\sigma_n^2 \to 0} 2(\underbrace{1 - \boldsymbol{\beta}_{\sim s}\mathbf{C}^{-1}\boldsymbol{\beta}_{\sim s}\theta_c\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}}}_{\lim=1\text{ (by II.11)}})\boldsymbol{\beta}'_b(\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\boldsymbol{\beta}_b\tau^*\sqrt{\frac{R}{\boldsymbol{\beta}'\boldsymbol{\beta}}}\varepsilon_{t+1} = 0.$$

The first term for the simple represents the systematic tendency for their intensity to decline when in power, as they respond to the overprediction of average outcomes. The ε_{t+1} term, for the simple and the complex, represents the effect of random shocks to y. Here, a negative shock reduces the intensity of the simple, as their belief in the effectiveness of the policies they deem relevant falls. Complex beliefs in these same policies also fall, but the complex belief in the efficacy of policies the simple deem irrelevant, and hence do not implement, rises, as the poor outcome under simple rule convinces the complex that these neglected policies are more effective than previously thought. These two effects offset each other, and complex intensity remains constant. In sum, a negative shock lowers the relative political intensity of the simple, hastening the transfer of power, with positive shocks having the opposite effect.

When the complex are in power asymptotically they implement policies $\boldsymbol{\beta}_c\sqrt{R/\boldsymbol{\beta}'\boldsymbol{\beta}}$ for the policies they believe are relevant and $\mathbf{0}_{k_{\sim c}x1}$ for those they believe are irrelevant, so the changes in intensity are seen to be

(II.13)
$$\lim_{\sigma_n^2 \to 0} \operatorname{plim}\ t(\bar{\boldsymbol\beta}'_{st+1}\bar{\boldsymbol\beta}_{st+1} - \bar{\boldsymbol\beta}'_{st}\bar{\boldsymbol\beta}_{st}) = \lim_{\sigma_n^2 \to 0} 2\tau^*\left[\boldsymbol\beta'_b\mathbf{A}^{-1}\boldsymbol\beta_b + \boldsymbol\beta'_{\sim c}\mathbf{0}_{k_{\sim c}x1}/\sigma_n^2\right]\sqrt{\frac{R}{\boldsymbol\beta'\boldsymbol\beta}}$$

$$*\left[\boldsymbol\beta'\sqrt{\frac{R}{\boldsymbol\beta'\boldsymbol\beta}}\boldsymbol\beta - \boldsymbol\beta'_s\sqrt{\frac{R}{\boldsymbol\beta'\boldsymbol\beta}}\tau^*\boldsymbol\beta_s + \varepsilon_{t+1}\right] = \frac{2\boldsymbol\beta'\boldsymbol\beta[\tau^*-1]}{\theta_s\tau^{*2}+\theta_c} + \sqrt{\frac{\boldsymbol\beta'\boldsymbol\beta}{R}}\frac{2\tau^*\varepsilon_{t+1}}{\theta_s\tau^{*2}+\theta_c},$$

$$\lim_{\sigma_n^2 \to 0} \operatorname{plim} t(\bar{\boldsymbol\beta}'_{ct+1}\bar{\boldsymbol\beta}_{ct+1} - \bar{\boldsymbol\beta}'_{ct}\bar{\boldsymbol\beta}_{ct}) = \lim_{\sigma_n^2 \to 0}\begin{bmatrix}\boldsymbol\beta'_b & \boldsymbol\beta'_{\sim s}\end{bmatrix}\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}' & \mathbf{C}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol\beta_b\\ \boldsymbol\beta_{\sim s}\end{bmatrix}\sqrt{\frac{R}{\boldsymbol\beta'\boldsymbol\beta}}\varepsilon_{t+1}$$

$$= \lim_{\sigma_n^2 \to 0} 2\left[\begin{array}{l}(1-\underbrace{(\boldsymbol\beta_{\sim s}\mathbf{C}^{-1}\boldsymbol\beta_{\sim s})\frac{\theta_c R}{\boldsymbol\beta'\boldsymbol\beta}}_{\text{lim=1 (by II.11)}})^2\boldsymbol\beta'_b(\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\boldsymbol\beta_b + \boldsymbol\beta_{\sim s}\mathbf{C}^{-1}\boldsymbol\beta_{\sim s}\\ +(\underbrace{\boldsymbol\beta'_{\sim s}\mathbf{C}^{-1}\mathbf{B}'}_{\text{lim}=\boldsymbol\beta'_b\text{ (by II.8 \& II.11)}} - \boldsymbol\beta'_b)(\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{C}^{-1}\boldsymbol\beta_{\sim s}\end{array}\right]\sqrt{\frac{R}{\boldsymbol\beta'\boldsymbol\beta}}\varepsilon_{t+1} = \sqrt{\frac{\boldsymbol\beta'\boldsymbol\beta}{R}}\frac{2\varepsilon_{t+1}}{\theta_c}.$$

Once again, the change in simple beliefs contains a systematic component, this time consisting of the gradual increase in bias and intensity as outcomes under the complex are consistently better than expected. Both simple and complex respond to the realization of the output shock ε, but the impact on the intensity of the complex is greater as, given that $\theta_c/\theta_s \to \tau^*$ as $\sigma_n^2 \to 0$, we have

(II.14)
$$\lim_{\sigma_n^2 \to 0}\frac{\tau^*}{\theta_s\tau^{*2}+\theta_c} = \frac{\tau^*}{\theta_c(\tau^*+1)} < \frac{1}{\theta_c}.$$

A negative shock reduces the belief in the effectiveness of policies of both types, but the effects on intensity are greater for the complex, for whom intensity depends upon a wider range of policies, all of which are seen to be failing. Consequently, negative shocks accelerate regime change, ushering in further negative outcomes as the simple implement misguidedly narrow and intense policies, while positive shocks lengthen the time the complex hold onto power and the polity continues to benefit from a full range of moderate policy actions. These results are those described in the paper and at the beginning of this appendix.

III: Results and Proofs on Berk-Nash Equilibria

We continue to use the notation and modelling framework established at the beginning of the paper's Appendix. We now define a Berk-Nash equilibrium adapted to our environment:

Definition 1: A Berk-Nash equilibrium consists of beliefs for $i \in \{S,C\}$ with mean $\overline{\boldsymbol{\beta}}_i$, a policy choice $\mathbf{x}_i$, and a probability that type S is in power, $\theta_s \in [0,1]$, such that:

(1a) *Optimal actions*: $\mathbf{x}_i$, is the optimal action given mean beliefs $\overline{\boldsymbol{\beta}}_i$ and so $\mathbf{x}_i = \mathbf{x}_i^*$.

(1b) *Power sharing according to intensity*: θ_s = 1 (0) if $\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s$ > (<) $\overline{\boldsymbol{\beta}}_c'\overline{\boldsymbol{\beta}}_c$; if $\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s = \overline{\boldsymbol{\beta}}_c'\overline{\boldsymbol{\beta}}_c$, $\theta_s \in [0,1]$.

(1c) *Beliefs minimize Kullback-Leibler distance*: Given actions $\mathbf{x}_c$, $\mathbf{x}_s$ and θ_s, each vector in the support of *i*'s beliefs solves, according to their subjective model:

$$\min_{\boldsymbol{\beta}_i} E_{\mathbf{n}} E_{\varepsilon}\left[\theta_s \ln \frac{f(\varepsilon)}{f(\boldsymbol{\beta}'(\mathbf{x}+\mathbf{n}) - \overline{\boldsymbol{\beta}}_i'(\mathbf{x}_{is}+\mathbf{n}_i)+\varepsilon)} + (1-\theta_s)\ln \frac{f(\varepsilon)}{f(\boldsymbol{\beta}'(\mathbf{x}+\mathbf{n}) - \overline{\boldsymbol{\beta}}_i'(\mathbf{x}_{ic}+\mathbf{n}_i)+\varepsilon)}\right]$$

Proposition III.1: For small enough σ_n^2 there exists a unique Berk-Nash equilibrium. In equilibrium, $\overline{\boldsymbol{\beta}}_c = \boldsymbol{\beta}_c$, $\overline{\boldsymbol{\beta}}_s = \tau^*\boldsymbol{\beta}_s$ and $\theta_s = (1-\tau^*\sigma_n^2/R)/(1+\tau^*)$.

Proof of Proposition III.1: First, it is easy to see that the unique KL minimiser for the belief of C is $\overline{\boldsymbol{\beta}}_c = \boldsymbol{\beta}_c$, as $\boldsymbol{\beta}_{\sim c}$ is a vector of 0s. Next, given the optimal policies, we solve for the minimiser of *S*'s KL condition. Taking the FOC of the term in (1c) with respect to $\overline{\boldsymbol{\beta}}_s$, we have:

$$\text{(III.1)} \qquad E_{\mathbf{n}} E_{\varepsilon}\left[\begin{array}{c} \theta_s^* \dfrac{f'(\boldsymbol{\beta}_s'\mathbf{x}_s^* + \boldsymbol{\beta}'\mathbf{n} - \overline{\boldsymbol{\beta}}_s'(\mathbf{x}_s^*+\mathbf{n}_s)+\varepsilon)}{f(\boldsymbol{\beta}_s'\mathbf{x}_s^* + \boldsymbol{\beta}'\mathbf{n} - \overline{\boldsymbol{\beta}}_s'(\mathbf{x}_s^*+\mathbf{n}_s)+\varepsilon)}(\mathbf{x}_s^*+\mathbf{n}_s) \\ +(1-\theta_s^*)\dfrac{f'(\boldsymbol{\beta}_c'\mathbf{x}_c^* + \boldsymbol{\beta}'\mathbf{n} - \overline{\boldsymbol{\beta}}_s'(\mathbf{x}_{sc}^*+\mathbf{n}_s)+\varepsilon)}{f(\boldsymbol{\beta}_c'\mathbf{x}_c^* + \boldsymbol{\beta}'\mathbf{n} - \overline{\boldsymbol{\beta}}_s'(\mathbf{x}_{sc}^*+\mathbf{n}_s)+\varepsilon)}(\mathbf{x}_{sc}^*+\mathbf{n}_s) \end{array}\right] = \mathbf{0}_{k_s x1}$$

$$\Rightarrow -E_{\mathbf{n}}\left[\begin{array}{c} \theta_s^*(\mathbf{x}_s^*+\mathbf{n}_s)[\boldsymbol{\beta}_s'\mathbf{x}_s^* + \boldsymbol{\beta}'\mathbf{n} - \overline{\boldsymbol{\beta}}_s'(\mathbf{x}_s^*+\mathbf{n}_s)] \\ +(1-\theta_s^*)(\mathbf{x}_{sc}^*+\mathbf{n}_s)[\boldsymbol{\beta}_c'\mathbf{x}_c^* + \boldsymbol{\beta}'\mathbf{n} - \overline{\boldsymbol{\beta}}_s'(\mathbf{x}_{sc}^*+\mathbf{n}_s)] \end{array}\right] = \mathbf{0}_{k_s x1}$$

$$\Rightarrow \theta_s^*\mathbf{x}_s^*[\boldsymbol{\beta}_s'\mathbf{x}_s^* - \overline{\boldsymbol{\beta}}_s'\mathbf{x}_s^*] + (1-\theta_s^*)\mathbf{x}_{sc}^*[\boldsymbol{\beta}_c'\mathbf{x}_c^* - \overline{\boldsymbol{\beta}}_s'\mathbf{x}_{sc}^*] = -(\boldsymbol{\beta}_s - \overline{\boldsymbol{\beta}}_s)\sigma_n^2,$$

where in moving to the second line we make use of the assumption that *f* is the density of a mean zero normal, and we follow the notation established earlier that a single subscript *i* denotes the policies deemed relevant by type *i* and a double subscript *ij* denotes these policies in periods of *j* rule, so that $\mathbf{x}_i^* = \mathbf{x}_{ii}^*$ are the optimal policies of type *i* but $\mathbf{x}_{ij}^*$ are the optimal policies of *j* deemed relevant by type *i*.

Arguments similar to those in Section II.A of the paper and continuity can be used to show that when σ_n^2 is sufficiently small in equilibrium θ_s^* is interior and so we have equal intensity,

(III.2) $$\overline{\boldsymbol{\beta}}_s'\overline{\boldsymbol{\beta}}_s = \overline{\boldsymbol{\beta}}_c'\overline{\boldsymbol{\beta}}_c = \boldsymbol{\beta}_c'\boldsymbol{\beta}_c = \boldsymbol{\beta}'\boldsymbol{\beta}.$$

Moreover, arguments similar to those in Section II.C can be used to show that optimal policies must be colinear when σ_n^2 is sufficiently small. Plugging the colinear optimal action and equal intensity into the FOC of the KL, we get the unique solution stated in the Proposition.