A Appendix: Details of Derivations

A.1 Worst-Case NRE Beliefs

Suppose that the policy commitment is of the conditionally linear form

$$\pi_{t+1} = p_t^0 + p_t^1 w_{t+1} \tag{A.1}$$

for some process $\{p_t^0(h_t, p_{-1}^0)\}$ and some deterministic sequence $\{p_t^1\}$. The problem of the "malevolent agent" in any state of the world at date t (corresponding to a history h_t up to that point) is to choose a function specifying m_{t+1} as a function of the realization of w_{t+1} so as to maximize

$$\frac{1}{2}[\pi_t^2 + \lambda(x_t - x^*)^2] - \theta E_t[m_{t+1} \log m_{t+1}] \tag{A.2}$$

subject to the constraint that $E_t m_{t+1} = 1$, where at each date x_t is implied by the equilibrium relation

$$\pi_t = \kappa x_t + \beta E_t[m_{t+1}\pi_{t+1}] + u_t. \tag{A.3}$$

It is obvious that the choice of the random variable m_{t+1}

matters only through its consequences for the relative entropy (which affects the objective (A.2)) on the one hand, and its consequences for PS expected inflation (which affects the constraint (A.3) on the other. Hence in the case of any $\theta > 0$, the worst-case beliefs will minimize the relative entropy $E_t[m_{t+1} \log m_{t+1}]$ subject to the constraints that

$$E_t m_{t+1} = 1, \qquad E_t[m_{t+1}\pi_{t+1}] = \bar{\pi}_t, \tag{A.4}$$

whatever degree of distortion the PS inflation expectation $\bar{\pi}_t$ may represent. I first consider this sub-problem.

Since $r(m) \equiv m \log m$ is a strictly convex function of m, such that $r'(m) \to -\infty$ as $m \to 0$ and $r'(m) \to +\infty$ as $m \to +\infty$, it is evident that there is a unique, interior optimum, in which the first-order condition

$$r'(m_{t+1}) = \phi_{1t} + \phi_{2t}\pi_{t+1}$$

holds in each state at date $t+1$, where ϕ_{1t}, ϕ_{2t} are Lagrange multipliers associated with the two constraints (A.4). This implies that

$$\log m_{t+1} = c_t + \phi_{2t}\pi_{t+1} \tag{A.5}$$

in each state, for some constant c_t. The two constants c_t and ϕ_{2t} in (A.5) are then the values that satisfy the two constraints (A.4).

Under the assumption of a conditionally linear policy (A.1), π_{t+1} is conditionally normally distributed, so that (A.5) implies that m_{t+1} is conditionally log-normal.[54] It follows that

$$\begin{aligned} \log E_t m_{t+1} &= E_t[\log m_{t+1}] + \frac{1}{2}\mathrm{var}_t[\log m_{t+1}] \\ &= c_t + \phi_{2t} p_t^0 + \frac{1}{2}\phi_{2t}^2 |p_t^1|^2. \end{aligned}$$

Hence the first constraint (A.4) is satisfied if and only if

$$c_t = -\phi_{2t} p_t^0 - \frac{1}{2}\phi_{2t}^2 |p_t^1|^2. \tag{A.6}$$

Under the worst-case beliefs, the PS perceives the conditional probability density for w_{t+1} to be $\tilde{f}(w_{t+1}) = m_{t+1}(w_{t+1})f(w_{t+1})$, where $f(\cdot)$ is the standard normal density. Hence

$$\begin{aligned} \log \tilde{f}(w) &= \log m_{t+1}(w) + \log f(w) \\ &= c_t + \phi_{2t}\pi_{t+1} - \frac{1}{2}log 2\pi - \frac{1}{2}w^2 \\ &= -\frac{1}{2}log 2\pi - \frac{1}{2}[w - \phi_{2t}p_t^1]^2, \end{aligned}$$

using (A.5) to substitute for m_{t+1} in the second line, and (A.1) and (A.6) to substitute for π_{t+1} and c_t respectively in the third line. But this is just the log density function for a variable that is distributed as $N(\mu_t, 1)$, where the bias in the perceived conditional expectation of w_{t+1} is $\mu_t = \phi_{2t} p_t^1$. Hence

$$\hat{E}_t \pi_{t+1} = p_t^0 + p_t^1 \mu_t = p_t^0 + \phi_{2t}|p_t^1|^2,$$

and the second constraint (A.4) is satisfied if and only if[55]

$$\phi_{2t} = \frac{\bar{\pi}_t - p_t^0}{|p_t^1|^2}. \tag{A.7}$$

[54] This is one of the main reasons for the convenience of restricting our attention to linear policies in this paper.

[55] Here I assume that $p_t^1 \neq 0$. If $p_t^1 = 0$, the constraint is satisfied regardless of the distortion chosen by the "malevolent agent," as long as $\bar{\pi}_t = p_t^0$, which is necessarily the case. In this case, c_t and ϕ_{2t} are not separately identified, but (A.6) suffices to show that $m_{t+1} = 1$ with certainty.

Condition (A.6) then uniquely determines c_t as well, and m_{t+1} is completely described by (A.5), once we have determined the value of $\bar{\pi}_t$ that should be chosen by the "malevolent agent." Note that the bias μ_t is given by expression (14), as asserted in the text.

The relative entropy of the worst-case beliefs will then be equal to

$$\begin{aligned} R_t^{pess} = \hat{E}_t[\log m_{t+1}] &= c_t + \phi_{2t}\hat{E}_t\pi_{t+1} \\ &= \frac{1}{2}\frac{(\bar{\pi}_t - p_t^0)^2}{|p_t^1|^2}, \end{aligned} \quad \text{(A.8)}$$

using (A.6) and (A.7). This is proportional to the squared distance between the PS inflation forecast and that of the central bank; but for any given size of gap between the two, the size of the distortion of probabilities that is required is smaller the larger is $|p_t^1|$.[56]

It remains to determine the worst-case choice of $\bar{\pi}_t$.[57] It follows from (A.3) that

$$(x_t^{pess} - x^*)^2 = \frac{1}{\kappa^2}(\pi_t - u_t - \kappa x^* - \beta\bar{\pi}_t)^2. \quad \text{(A.9)}$$

Substituting this for the squared output gap and (A.8) for the relative entropy in (A.2), we obtain an objective for the "malevolent agent" that is a quadratic function $Q(\bar{\pi}_t; u_t, \pi_t, p_t)$ of the distorted inflation forecast $\bar{\pi}_t$, and otherwise independent of the distorted beliefs; thus $\bar{\pi}_t$ is chosen to maximize this function. The function is strictly concave (because the coefficient multiplying $\bar{\pi}_t^2$ is negative) if and only if p_t^1 satisfies the inequality

$$|p_t^1|^2 < \frac{\theta}{\beta^2}\frac{\kappa^2}{\lambda}. \quad \text{(A.10)}$$

If the inequality is reversed, the function Q is instead *convex*, and is minimized rather than maximized at the value of $\bar{\pi}_t$ that satisfies the first-order condition $Q_{\bar{\pi}} = 0$. But in this case, the "malevolent agent" can achieve an unboundedly large positive value of the objective (A.2), as stated in the text; and a robustly optimal policy can never involve a value of p_t^1 this large.

In the case that (A.10) holds with equality, Q is linear in $\bar{\pi}_t$, and it is again possible for the "malevolent agent" to achieve an unboundedly large positive value of

[56] Equation (A.8) again assumes that $p_t^1 \neq 0$. In the event that $p_t^1 = 0$, it follows from the previous footnote that the relative entropy of the worst-case beliefs will equal zero.

[57] The analysis here assumes that $p_t^1 \neq 0$. If $p_t^1 = 0$, there is no choice about the value of $\bar{\pi}_t$; it must equal p_t^0.

the objective through an extreme choice of $\bar{\pi}_t$, except in the special case that

$$p_t^0 = \beta^{-1}(\pi_t - u_t - \kappa x^*), \tag{A.11}$$

so that the linear function has a slope of exactly zero. Thus unless p_t^0 satisfies (A.11), p_t^1 must satisfy the bound (A.10) in order for the objective (A.2) to have a finite maximum. Even in the special case that (A.11) holds exactly, p_t^1 must satisfy a variant of (A.10) in which the strict inequality is replaced by a weak inequality.

When (A.10) holds, the maximum value of Q occurs for the value of $\bar{\pi}_t$ such that $Q_{\bar{\pi}} = 0$. This implies that the worst-case value of $\bar{\pi}_t$ is

$$\bar{\pi}_t = \Delta_t^{-1}\left[p_t^0 - (\pi_t - u_t - \kappa x^*)\frac{\beta\lambda}{\theta\kappa^2}|p_t^1|^2\right], \tag{A.12}$$

$$\Delta_t \equiv 1 - \frac{\beta^2}{\theta}\frac{\lambda}{\kappa^2}|p_t^1|^2 > 0, \tag{A.13}$$

as stated in the text. Substituting this solution into (A.8) and (A.9), one obtains the implied output gap (17) and and relative entropy (18) under the worst-case NRE beliefs, as stated in the text. Substituting these expressions into the objective (A.2), one obtains an objective for the CB of the form

$$\hat{\mathcal{L}}(p; p_{-1}^1, \rho) = \mathrm{E}\sum_{t=0}^{\infty}\beta^t L(p_{t-1}; p_t; w_t), \tag{A.14}$$

in which the period loss is given by

$$L(p_{t-1}; p_t; w_t) \equiv \frac{1}{2}\pi_t^2 + \frac{\lambda}{2\kappa^2\Delta_t}[\pi_t - u_t - \kappa x^* - \beta p_t^0]^2, \tag{A.15}$$

where $0 < \Delta_t < 1$ is the function of p_t^1 defined by (A.13), π_t is the function of p_{t-1} and w_t defined by (A.1), and $u_t = \sigma_u w_t$. Note that we can alternatively write

$$L(p_{t-1}; p_t; w_t) = \tilde{L}(\pi_t; p_t; w_t),$$

where the function $\tilde{L}$ is defined by the right-hand side of (A.15), since the coefficients p_{t-1} only enter through their consequences for the value of π_t.[58]

When, instead, (A.10) holds with equality, and (A.11) holds as well, the worst-case value of $\bar{\pi}_t$ is indeterminate, but the maximized value of (A.2) is nonetheless well-defined, and equal to zero. In this case, the period loss function is equal to

$$L(p_{t-1}; p_t; w_t) = \frac{1}{2}\pi_t^2.$$

[58] This alternative expression for the period loss function is convenient in Appendix A.3.

When neither this case nor the one discussed in the previous paragraph applies, we can define $L(p_{t-1}; p_t; s_t)$ as being equal to $+\infty$. The function is then defined (but possibly equal to $+\infty$) for all possible values of its arguments.

Note also that $L(p_{t-1}; p_t; s_t)$ is necessarily non-negative, since for any values of the arguments, it is possible for the "malevolent agent" to obtain a non-negative value of (A.2) by choosing $m_{t+1} = 1$ in all states; the maximized value of (A.2) is then necessarily at least this high. It follows that both the conditional expectations and the infinite sum in (A.14) are sums (or integrals) of non-negative quantities; hence both are well-defined (though possibly equal to $+\infty$) for all possible values of the arguments. Thus the CB objective (A.14) is well-defined for an arbitrary conditionally-linear policy $\{p_t\}$ and arbitrary initial conditions (p^1_{-1}, ρ).

A.2 Robustly Optimal Linear Policy

Given the worst-case PS beliefs characterized in the previous appendix, the problem of the CB is to choose a $\{p_t\}$ for all $t \geq 0$ so as to minimize (A.14), for given initial conditions p^1_{-1} and a distribution ρ of possible values for p^0_{-1}. The CB must choose a policy under which p^0_t may depend on both p^0_{-1} and the history of shocks h_t, but p^1_t must be a deterministic function of time.

One can show that the objective (A.14) is a convex function of the sequence $\{p_t\}$. I begin by noting that (A.2) is a convex function of π_t and x_t, for any choice of $m_{t+1}(\cdot)$. Then since (A.3) is a linear relation among π_t, x_t, and $\pi_{t+1}(\cdot)$, it follows that, taking as given the choice of $m_{t+1}(\cdot)$, the value of (A.2) implied by any choice of $\pi_{t+1}(\cdot)$ by the CB is a convex function of π_t and $\pi_{t+1}(\cdot)$. Similarly, since (A.1) is linear, the value of (A.2) implied by any choice of p_t is a convex function of p_{t-1} and p_t, for any choice of $m_{t+1}(\cdot)$. Then since the maximum of a set of convex functions is a convex function, it follows that the maximized value of (A.2) is also a convex function of p_{t-1} and p_t. Thus $L(p_{t-1}; p_t; w_t)$ is a convex function of (p_{t-1}, p_t). Finally, a sum of convex functions is convex; this implies that (A.14) is a convex function of the sequence $\{p_t\}$.

Convexity implies that the CB's optimal policy can be characterized by a system of first-order conditions,[59] according to which

$$L_3(p_{t-1}; p_t; w_t) + \beta \mathrm{E}_t L_1(p_t; p_{t+1}; w_{t+1}) = 0 \tag{A.16}$$

for each possible history h_t at any date $t \geq 0$, and

$$\mathrm{E}[L_4(p_{t-1}; p_t; w_t) + \beta L_2(p_t; p_{t+1}; w_{t+1})] = 0 \tag{A.17}$$

for each date $t \geq 0$. Here L_1 through L_4 denote the partial derivatives of $L(p^0_{t-1}, p^1_{t-1}; p^0_t, p^1_t; w_t)$ with respect to its first through fourth arguments, respectively. Condition (A.16) is the first-order condition for the optimal choice of $p^0_t(h_t, p^0_{-1}$, and (A.17) is the corresponding condition for the optimal choice of p^1_t (which must take the same value in all states of the world at date t).

Note that it follows from the characterization in the previous appendix that for any plan satisfying (A.10), the partial derivatives just referred to are well-defined, and equal to

$$L_1(p_{t-1}; p_t; w_t) = \pi_t + \frac{\lambda}{\kappa^2} \frac{\pi_t - u_t - \kappa x^* - \beta p^0_t}{\Delta_t},$$

$$L_2(p_{t-1}; p_t; w_t) = L_1(p_{t-1}; p_t; w_t) w_t,$$

$$L_3(p_{t-1}; p_t; w_t) = -\beta \frac{\lambda}{\kappa^2} \frac{\pi_t - u_t - \kappa x^* - \beta p^0_t}{\Delta_t},$$

$$L_4(p_{t-1}; p_t; w_t) = \frac{\beta^2}{\theta} \left(\frac{\lambda}{\kappa^2} \right)^2 \left(\frac{\pi_t - u_t - \kappa x^* - \beta p^0_t}{\Delta_t} \right)^2 p^1_t.$$

Substituting (A.1) for π_t and (A.13) for Δ_t in these expressions, one can express the first-order conditions (A.16) – (A.17) as restrictions upon the sequence $\{p_t\}$.

Taking as given the deterministic sequence $\{p^1_t\}$, one observes that (A.16) is a linear stochastic difference equation for the evolution of the process $\{p^0_t\}$, with coefficients that are time-varying insofar as they involve the coefficients $\{p^1_t\}$. One can show that these linear equations must have a linear solution of the form (20). Here there is no need to give a general expression for the coefficients of this solution, as

[59] This sequence of first-order conditions by itself is necessary but not sufficient for an optimum; in order to prove that a solution to the FOCs represents an optimum, one must also verify a transversality condition. Here, however, we are interested only in the steady-state solution to the FOCs, which necessarily satisfies the transversality condition. Hence the steady-state solution characterized below does represent the policy that minimizes (A.14) under the self-consistent specification of the initial conditions.

we are interested only in the existence of a steady state. In such a steady state, if it exists, p_t^1 is equal to some constant value $\bar{p}^1$ for all t; so it suffices to consider the solution to (A.16) in this case.

Under the assumption that $p_t^1 = \bar{p}^1$ for all $t \geq -1$, (A.16) is a stochastic linear difference equation for the process $\{p_t^0\}$ of the form

$$E_t[A(L)p_{t+1}^0] = (\sigma_u - \bar{p}^1)w_t, \tag{A.18}$$

where

$$A(L) \equiv \beta - \left(1 + \beta + \frac{\kappa^2\bar{\Delta}}{\lambda}\right) L + L^2.$$

(Here $\bar{\Delta}$ is the constant value of Δ_t implied by the constant value $\bar{p}^1$.) By factoring the lag polynomial in (A.18), one can easily show that (A.18) has a unique stationary solution,[60] given by

$$p_t^0 = \mu p_{t-1}^0 - \mu(\sigma_u - \bar{p}^1)w_t, \tag{A.19}$$

where $0 < \mu < 1$ is the smaller root of the characteristic equation (24) given in the text. Note that a stationary solution exists regardless of the value assumed for $\bar{p}^1$, as long as it satisfies (A.10), for the quadratic equation is easily seen to have a root in that interval in the case of any $\bar{\Delta} > 0$. In fact, since $0 < \bar{\Delta} < 1$, one can show that $\mu^{RE} < \mu < 1$, where μ^{RE} is the root in the RE case (corresponding to $\bar{\Delta} = 1$).

The law of motion (A.19) implies that if the unconditional distribution for p_{t-1}^0 is $N(\mu_{p,t-1}, \sigma_{p,t-1}^2)$, then (given the assumption that w_t is i.i.d. $N(0,1)$) the unconditional distribution for p_t^0 is also normal, with mean and variance

$$\mu_{p,t} = \mu\mu_{p,t-1}, \qquad \sigma_{p,t}^2 = \mu^2[\sigma_{p,t-1}^2 + (\sigma_u - \bar{p}^1)^2].$$

These difference equations have a unique fixed point, corresponding to the stationary or ergodic distribution implied by the law of motion (A.19), namely,

$$\bar{\mu}_p = 0, \qquad \bar{\sigma}_p^2 = \frac{\mu^2(\sigma_u - \bar{p}^1)^2}{1 - \mu^2}.$$

I turn next to the implications of conditions (A.17). Note that for each period $t \geq 0$, the left-hand side of this equation involves the values of the three quantities $(p_{t-1}^1, p_t^1, p_{t+1}^1)$ and the unconditional joint distribution of $(p_{t-1}^0, p_t^0, p_{t+1}^0; w_t, w_{t+1})$.

[60] This is the solution to the FOCs that satisfies the transversality condition and hence that corresponds to the process that minimizes (A.14). For a generalization of this characterization to the case in which the process $\{u_t\}$ follows a more general linear process, see Woodford (2005, Appendix A.2).

Given the assumption of a normal distribution $N(\mu_{p,t-1}, \sigma^2_{p,t-1})$ for p^0_{t-1} and the law of motion (20) for $\{p^0_t\}$ under optimal policy, we can write this joint distribution as a function of the parameters $(\mu_{p,t-1}, \sigma^2_{p,t-1})$ of the marginal distribution for p^0_{t-1} and the parameters (ψ_t, ψ_{t+1}) of the conditional distribution $(p^0_t, p^0_{t+1}; w_t, w_{t+1}|p^0_{t-1})$. (Recall that ψ_t denotes the vector of coefficients of the law of motion (20).) Hence the left-hand side of (A.17) is a function of the form

$$g(p^1_{t-1}, p^1_t, p^1_{t+1}; \mu_{p,t-1}, \sigma^2_{p,t-1}; \psi_t, \psi_{t+1}),$$

as asserted in (22). Once again, we need not further discuss the form of this equation except in the case of a steady-state solution.

Using the solution above for the unconditional joint distribution of $(p^0_{t-1}, p^0_t, p^0_{t+1}; w_t, w_{t+1})$ in the case of self-consistent initial conditions, condition (A.17) then becomes a second-order nonlinear difference equation in p^1_t (the coefficients of which depend, however, on the assumed value of $\bar{p}^1$). One observes that

$$\begin{aligned} \mathrm{E}[L_4(p_{t-1}; p_t; w_t)] &= \frac{\beta^2}{\theta}\left(\frac{\lambda}{\kappa^2}\right)^2 \frac{\bar{p}^1}{\bar{\Delta}^2}\mathrm{E}[(\pi_t - u_t - \kappa x^* - \beta p^0_t)^2] \\ &= \frac{\beta^2}{\theta}\left(\frac{\lambda}{\kappa^2}\right)^2 \frac{\bar{p}^1}{\bar{\Delta}^2}[a + 2b\bar{p}^1 + (\bar{p}^1)^2], \end{aligned}$$

where

$$a \equiv \mathrm{E}[(p^0_{t-1} - u_t - \kappa x^* - \beta p^0_t)^2],$$

$$b \equiv \mathrm{E}[w_t(p^0_t - u_t - \kappa x^* - \beta p^0_t)].$$

Similarly, one can show that

$$\begin{aligned} \mathrm{E}[L_2(p_t; p_{t+1}; w_{t+1})] &= \mathrm{E}[\pi_{t+1} w_{t+1}] + \frac{\lambda}{\kappa^2\bar{\Delta}}\mathrm{E}[(\pi_{t+1} - u_{t+1} - \kappa x^* - \beta p^0_{t+1})w_{t+1}] \\ &= \bar{p}^1 + \frac{\lambda}{\kappa^2\bar{\Delta}}[\bar{p}^1 + b]. \end{aligned}$$

Hence condition (A.17) is equivalent to

$$f(\bar{p}^1) \equiv \frac{\beta^2}{\theta}\left(\frac{\lambda}{\kappa^2}\right)^2 \frac{c}{\bar{\Delta}^2}\bar{p}^1 + \bar{p}^1 + \frac{\lambda}{\kappa^2\bar{\Delta}}[\bar{p}^1 + b] = 0, \tag{A.20}$$

where

$$c \equiv a + 2b\bar{p}^1 + (\bar{p}^1)^2.$$

A robustly optimal linear policy then exists if and only if (A.20) has a solution $\bar{p}^1$ that satisfies the bound (A.10). Of course, in defining the function $f(\cdot)$, one must take account of the dependence of c and $\bar{\Delta}$ on the value of $\bar{p}^1$.

When $\{p_t^0\}$ evolves in accordance with the stationary dynamics (A.19), the above definitions imply that

$$\begin{aligned} a &= (\kappa x^*)^2 + \mathrm{E}\{[(1-\beta\mu)p_{t-1}^0 - (\sigma_u - \beta\mu(\sigma_u - \bar{p}^1))w_t]^2\} \\ &= (\kappa x^*)^2 + \frac{(1-\beta\mu)^2\mu^2}{1-\mu^2}(\sigma_u - \bar{p}^1)^2 + [(1-\beta\mu)\sigma_u + \beta\mu\bar{p}^1]^2, \end{aligned}$$

$$\begin{aligned} b &= -\sigma_u - \beta \mathrm{E}[p_t^0 w_t] \\ &= -(1-\beta\mu)\sigma_u - \beta\mu\bar{p}^1. \end{aligned}$$

I furthermore observe that $a = a_0 + b^2$, where

$$a_0 \equiv (\kappa x^*)^2 + \frac{(1-\beta\mu)^2\mu^2}{1-\mu^2}(\sigma_u - \bar{p}^1)^2 > 0.$$

Hence

$$c = a_0 + (b + \bar{p}^1)^2 > 0$$

can be signed for all admissible values of $\bar{p}^1$. Substituting this function of $\bar{p}^1$ for c and (A.13) for $\bar{\Delta}$ in (A.20) yields a nonlinear equation in $\bar{p}^1$, that is solved numerically in order to produce Figure 1.

One can easily show that a solution to this equation in the admissible range must exist. Note first that (A.10) can alternatively be written in the form

$$|\bar{p}^1| < \hat{p}^1 \equiv \frac{\kappa}{\lambda^{1/2}} \frac{\theta^{1/2}}{\beta}. \qquad \text{(A.21)}$$

I next observe that

$$f(0) = \frac{\lambda}{\kappa^2 \bar{\Delta}} b = -\frac{\lambda}{\kappa^2}(1-\beta\mu)\sigma_u < 0.$$

On the other hand, in the case of any finite θ, as $p^1 \to \hat{p}^1$, the first term in the expression (A.20) becomes larger than the other two terms, so that $f(p^1) > 0$ for any value of p^1 close enough to (while still below) the bound, Since the function $f(\cdot)$ is well-defined and continuous on the entire interval $[0,\ \hat{p}^1)$, there must be an intermediate value $0 < \bar{p}^1 < \hat{p}^1$ at which $f(\bar{p}^1) = 0$. Such a value satisfies both (A.10) and (A.20), and so describes a robustly optimal linear policy.

One can further establish that

$$0 < \bar{p}^1 < \mu\sigma_u, \qquad \text{(A.22)}$$

as asserted in the text. When evaluated at the value $p^1 = \mu\sigma_u$, the second two terms in (A.20) are equal to

$$-\frac{\lambda}{\kappa^2\bar{\Delta}} P(\mu)\sigma_u = 0,$$

where $P(\mu)$ is the polynomial defined in (24). Moreover, in the limiting case in which $\theta \to \infty$ (the RE case), the first term in condition (A.20) is identically zero, so that $f(\mu\sigma_u) = 0$, and $\bar{p}^1 = \mu\sigma_u$ is a solution.[61] Instead, when θ is finite, the first term is necessarily positive, so that $f(\mu\sigma_u) > 0$. If $\mu\sigma_u < \hat{p}^1$, this implies that there exists a solution to (A.17) such that (A.22) holds. If instead $\hat{p}^1 \leq \mu\sigma_u$, then (A.22) follows from the result in the previous paragraph. Hence in either case, the robustly optimal policy satisfies (A.22) for any finite θ, while the upper bound holds with equality in the limiting case of infinite θ.

Substitution of the law of motion (A.19) for p_t^0 in (A.12) leads to the solution

$$\bar{\pi}_t = \Lambda p_t^0 + \beta^{-1}(\bar{\Delta}^{-1} - 1)\kappa x^*,$$

where

$$\Lambda \equiv \bar{\Delta}^{-1} - \beta^{-1}\mu^{-1}(\bar{\Delta}^{-1} - 1).$$

Note that

$$\Lambda - 1 = (1 - \beta^{-1}\mu^{-1})(\bar{\Delta}^{-1} - 1) < 0,$$

since $0 < \beta, \mu, \bar{\Delta} < 1$, from which it follows that $\Lambda < 1$.

A.3 Existence and Stability of Robust Linear MPE

A robust linear MPE corresponds to a pair $(\bar{p}^1, \bar{\Delta})$ that satisfy equations

$$\bar{p}^1 = \frac{\lambda}{\kappa^2\bar{\Delta} + \lambda}\sigma_u > 0, \qquad \text{(A.23)}$$

$$\bar{\Delta} = 1 - \frac{\beta^2}{\theta}\frac{\lambda}{\kappa^2}|\bar{p}^1|^2, \qquad \text{(A.24)}$$

[61] It is easily seen to be the unique solution, since $f(p)$ is linear in this case. One can also show that this is the optimal policy without restricting attention to linear policies, as is done here; see Clarida *et al.* (1999) or Woodford (2003, chap. 7).

with $\bar{\Delta} > 0$ so that (A.10) is satisfied. Equivalently, we are looking for solutions to the two equations in the interval $0 < \bar{p}^1 < \hat{p}^1$, where $\hat{p}^1$ is defined by (A.21).

If we write these equations as $\bar{\Delta} = \Delta_1(\bar{p}^1)$ and $\bar{\Delta} = \Delta_2(\bar{p}^1)$ respectively, we observe that $\Delta_1(p)$ is a decreasing, strictly concave function for all $p > 0$, while $\Delta_2(p)$ is a decreasing, strictly convex function over the same domain. Moreover, $\Delta_1(p) < \Delta_2(p)$ for all small enough $p > 0$ (as $\Delta_2(p) \to +\infty$ as $p \to 0$), and also for all large enough p (as $\Delta_1(p) \to -\infty$ as $p \to +\infty$). Hence there are either *no* intersections of the two curves with $\bar{p}^1 > 0$, or *two* intersections, or a single intersection at a point of tangency between the two curves.

The slopes of the two curves are furthermore given by

$$\Delta_1'(p) = -2\frac{\beta^2}{\theta}\frac{\lambda}{\kappa^2}p,$$

$$\Delta_2'(p) = -\frac{\lambda}{\kappa^2}\frac{\sigma_u}{p^2}.$$

From these expressions one observes that $\Delta_2'(p)$ is less than, equal to, or greater than $\Delta_1'(p)$ according to whether p is less than, equal to, or greater than $\tilde{p}^1$, where

$$\tilde{p}^1 \equiv \left(\frac{\theta}{\beta^2}\frac{\sigma_u}{2}\right)^{1/3} > 0.$$

From this it follows that there are two intersections if and only if $\Delta_2(\tilde{p}^1) < \Delta_1(\tilde{p}^1)$, which holds if and only if $\sigma_u < \sigma_u^*$, where σ_u^* is defined as in (37).[62] Similarly, the two curves are tangent to each other if and only if $\sigma_u = \sigma_u^*$; in this case, the unique intersection is at $\bar{p}^1 = \tilde{p}^1$. And finally, the two curves fail to intersect if and only if $\sigma_u > \sigma_u^*$.

It remains to consider how many of these intersections occur in the interval $0 < \bar{p}^1 < \hat{p}^1$. One notes that there is exactly one solution in that interval (and hence a unique robust linear MPE) if and only if $\Delta_2(\hat{p}^1) < 0$, which holds if and only if $\sigma_u < \hat{p}^1$. When $\sigma_u = \hat{p}^1$ exactly, $\Delta_2(\hat{p}^1) = \Delta_1(\hat{p}^1) = 0$, and the curves intersect at $\bar{p}^1 = \hat{p}^1$. This is the larger of two solutions for $\bar{p}^1$ if and only if

$$\Delta_1'(\hat{p}^1) < \Delta_2'(\hat{p}^1), \tag{A.25}$$

which holds if and only if $\lambda\kappa^2 < 2$. In this case, as σ_u is increased further, the larger of the two solutions for $\bar{p}^1$ decreases with σ_u, so that there are two solutions in the

[62] It is useful to note that this definition implies that $\sigma_u^* \geq \hat{p}^1$, with equality only if $\kappa^2/\lambda = 1/2$.

interval $(0, \hat{p}^1)$, until $\sigma_u = \sigma^*$, and the two solutions collapse into one, as the curves are tangent. (Note that $\sigma_u^* > \hat{p}^1$.) For still larger values of σ_u, there is no intersection, as explained in the previous paragraph.

If instead, $\lambda/\kappa^2 = 2$ exactly, then the curves are tangent when $\sigma_u = \hat{p}^1$ (which in this case is also equal to σ_u^*). At this point the only intersection occurs at $\bar{p}_1 = \hat{p}^1$ (which fails to satisfy condition (A.10)), and for larger values of σ_u there are no intersections. Finally, if $\lambda/\kappa^2 > 2$, then the inequality in (A.25) is reversed, and when $\sigma_u = \hat{p}^1$, the intersection at $\bar{p}^1 = \hat{p}^1$ is the smaller of the two solutions. (The smaller solution approaches $\hat{p}^1$ from below as σ_u increases to $\hat{p}^1$.) In this case, there are no solutions $\bar{p}^1 < \hat{p}^1$ when $\sigma_u = \hat{p}^1$. As σ_u increases further, the smaller solution continues to increase with σ_u, so that even for values of σ_u that continue to be less than or equal to σ_u^* (so that the curves continue to intersect), there are no solutions with $\bar{p}^1 < \hat{p}^1$. And for still larger values of σ_u, there are again no solutions at all. Hence in each case, the number of solutions is as described in the text.

The "expectational stability" analysis proposed in the text involves the properties of the map

$$\Phi(p) \equiv \Delta_2^{-1}(\Delta_1(p)).$$

Formally, a fixed point $\bar{p}^1$ of Φ (which corresponds to an intersection of the two curves studied above) is expectationally stable if and only if there exists a neighborhood P of $\bar{p}^1$ such that

$$\lim_{n\to\infty} \Phi^n(p) = \bar{p}^1$$

for any $p \in P$. Our observations above about the functions $\Delta_1(\cdot), \Delta_2(\cdot)$ imply that $\Phi(\cdot)$ is a monotonically increasing function. Hence a fixed point $\bar{p}^1$ is stable if and only if $\Phi'(\bar{p}^1) < 1$.

The above definition implies that

$$\Phi'(p) = \frac{\Delta_1'(p)}{\Delta_2'(\Delta_2^{-1}(\Delta_1(p)))} > 0.$$

Evaluated at a fixed point of Φ, this reduces to

$$\Phi'(\bar{p}^1) = \frac{\Delta_1'(\bar{p}^1)}{\Delta_2'(\bar{p}^1)}.$$

Hence the stability condition is satisfied if and only if

$$\Delta_2'(\bar{p}^1) < \Delta_1'(\bar{p}^1) < 0. \tag{A.26}$$

Because of the concavity of $\Delta_1(\cdot)$ and the convexity of $\Delta_2(\cdot)$, this condition necessarily holds at the fixed point with the smaller value of $\bar{p}^1$, and not at the higher value. Hence in Figure 3, it is the upper (dashed) branch of solutions that is expectationally unstable, while the lower (solid) branch of solutions is stable. We therefore conclude that regardless of the other parameter values, there is exactly one expectationally stable robust linear MPE for all values of σ_u below some positive critical value, and no robust linear MPE for values of σ_u greater than or equal to that value.

Finally, let us consider the way in which $\bar{p}^1$ changes as θ is reduced (indicating that a broader range of NRE beliefs are considered possible). Letting $\bar{p}^1$ be implicitly defined by the equation

$$\Delta_1(\bar{p}^1) = \Delta_2(\bar{p}^1),$$

the implicit function theorem implies that

$$\frac{d\bar{p}^1}{d\theta} = -\frac{\partial\Delta_1/\partial\theta}{\Delta_1' - \Delta_2'}. \tag{A.27}$$

It follows from (A.26) that in the case of an expectationally stable MPE, the denominator of the fraction in (A.27) is positive. We also observe that

$$\frac{\partial\Delta_1}{\partial\theta} = \frac{\beta^2}{\theta^2}\frac{\lambda}{\kappa^2}(\bar{p}^1)^2 > 0,$$

so that the numerator is positive as well, and hence $\bar{p}^1$ decreases as θ increases. This means that $\bar{p}^1$ increases as the CB's concern for robustness increases (corresponding to a lower value of θ, up until the point where there ceases to any longer be a robust linear MPE at all. In that case, as discussed in the text, we can think of the equilibrium sensitivity of inflation to cost-push shocks as being *unbounded;* so the conclusion that greater concern for robustness leads to greater sensitivity of inflation to cost-push shocks extends, in a looser sense, to that case as well.