A Note on

Optimal Monetary Policy

in the standard 3-equation NK-model

Ulrich Roschitsch*

May 2023 (last revised March 1, 2024)

`Link to most recent version`

Notation

Scalars: x. Vectors: $\boldsymbol{x}$. Matrices: $\underline{\boldsymbol{x}}$. Element i, j of matrix $\underline{\boldsymbol{x}}$: $[\underline{\boldsymbol{x}}]_{i,j} \in \mathbb{R}$. The converse, a matrix composed of the elements of the doubly array $\mathbb{N}^2 \supset I \times J \to \mathbb{R} : (i, j) \mapsto a_{i,j}$, is denoted $[a_{i,j}]_{i \in I,\, i \in J}$ (with the understanding that the dimension is $|I| \times |J|$). The Jacobian of $\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^n$ with respect to $\boldsymbol{x}$ at the point $\boldsymbol{y}$ is denoted $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}'}\big|_{\boldsymbol{y}}$. Terms of order $k \in \mathbb{R}$ in some variable are generically denoted $O(k)$ without explicit mention of the variable; this is typically some vector of exogenous (random) variables, $\boldsymbol{\xi}$, so that we have $O(k) := O(\|\boldsymbol{\xi}\|^k)$.

Main Text

Relative to the model from the lecture we make two changes: (i) We consider the case where the marginal product of labor in producing consumption goods may be decreasing; (ii) we consider price stickiness à la Rotemberg – firms must pay a cost for changing their price that which is quadratic in the size of the price adjustment. None of these changes affects the generality of the results.

One straightforward way to introduce a (weakly) decreasing marginal product of labor into our otherwise unaltered production side à la Rotemberg is to introduce "wholesale firms". The representative wholesale producer hires labor, n_t, at nominal wage W_t and produces $\xi_{a,t} n_t^{\alpha}$, $\alpha \in (0, 1]$ units of wholesale goods. Wholesale goods, y_t^w, are sold to

*University of Mannheim, Email: ulrich.roschitsch@uni-mannheim.de

variety producers at nominal price M_t who turn them into varieties using a linear technology. The FOC of wholesale producers is then:

$$M_t = \xi_{a,t}^{-1} W_t \alpha^{-1} n_t^{1-\alpha}.$$

The program of a variety producer j is then:

$$\max_{(P_t(j))_{t\geq 0}} \mathbb{E}_0 \sum_{t\geq 0} \beta^t \frac{\Lambda_t}{P_t} \left[P_t(j) y_t(j) - (1 - \tau^\ell \xi_{\tau,t}) M_t y_t^w(j) - P_t \frac{\kappa}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - 1 \right)^2 y_t \right] \text{ with}$$

$$y_t(j) = y_t^w(j) \wedge y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} y_t.$$

Notice that here the production subsidy shifter $\xi_{\tau,t}$ acts as an *inverse cost-push shock*! The linearized Phillips-curve then is:

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \frac{\epsilon - 1}{\kappa} (\widehat{w}_t + (1-\alpha)\widehat{n}_t - \widehat{\xi}_{a,t}) - \frac{1}{\kappa} \widehat{\xi}_{\tau,t}.$$

Up to a rule determining i_t, the equilibrium is summarized by:

$$\begin{aligned}
&(n) \quad \varphi \widehat{n}_t + \sigma \widehat{c}_t - \widehat{\xi}_{c,t} = \widehat{w}_t \\
&(b) \quad \widehat{c}_t = \mathbb{E}_t \widehat{c}_{t+1} - \frac{1}{\sigma}(i_t - \mathbb{E}_t \pi_{t+1} + \ln \beta + \mathbb{E}_t \Delta \widehat{\xi}_{c,t+1}) \\
&(PC) \quad \pi_t = \beta \mathbb{E}_t \pi_{t+1} + \frac{\epsilon - 1}{\kappa} (\widehat{w}_t + (1-\alpha)\widehat{n}_t - \widehat{\xi}_{a,t}) - \frac{1}{\kappa} \widehat{\xi}_{\tau,t} \\
&(y) \quad \widehat{y}_t = \widehat{\xi}_{a,t} + \alpha \widehat{n}_t \\
&(MC) \quad \widehat{y}_t = \widehat{c}_t
\end{aligned}$$

1 Efficient Allocation

The efficient allocation obtains in the absence of price stickiness and cost-push-shocks that drive a wedge between marginal costs and prices charged by monopolistically competitive variety producers. That is, we consider the allocation arising under $\kappa \to 0$, $\xi_\tau \equiv 0$. The system then reads

$$\begin{aligned}
&(n)^e \quad \varphi \widehat{n}_t^e + \sigma \widehat{y}_t^e - \widehat{\xi}_{c,t} = \widehat{w}_t^e \\
&(b)^e \quad \widehat{y}_t^e = \mathbb{E}_t \widehat{y}_{t+1}^e - \frac{1}{\sigma}(r_t^e + \mathbb{E}_t \Delta \widehat{\xi}_{c,t+1}) \\
&(PC)^e \quad \widehat{w}_t^e = \widehat{\xi}_{a,t} - (1-\alpha)\widehat{n}_t^e \\
&(y)^e \quad \widehat{y}_t^e = \widehat{\xi}_{a,t} + \alpha \widehat{n}_t^e
\end{aligned}$$

where r_t^e is defined as the real interest rate that makes $(b)^e$ hold. (Since prices are no longer sticky monetary policy is neutral and inflation simply adjusts such that $(b)^e$ holds; this is captured by simply introducing the efficient real rate, r_t^e.)

Now substitute $(PC)^e$ into $(n)^e$ and solve for n_t^e:

$$\widehat{n}_t^e = (1+\varphi-\alpha)^{-1}(\widehat{\xi}_{a,t} + \widehat{\xi}_{c,t} - \sigma \widehat{y}_t^e).$$

This we can plug into $(y)^e$ to find:

$$\boxed{\widehat{y}_t^e = \frac{1+\varphi}{1+\varphi+\alpha(\sigma-1)}\widehat{\xi}_{a,t} + \frac{\alpha}{1+\varphi+\alpha(\sigma-1)}\widehat{\xi}_{c,t}.}$$

This completely determines the efficient allocation as an explicit expression of exogeneous processes.

2 Allocation Gaps

In anticipation of the fact that the loss function will be expressed in deviations of the market allocation from the efficient allocation, we can use our just-derived results to rewrite the market equilibrium conditions to be in terms of these deviations. Specifically, define $\widetilde{y}_t := \widehat{y}_t - \widehat{y}_t^e$ as the output gap. The first step to finding the law of motion of the gap variables is to rewrite the Phillips-curve. First, insert $\widehat{n}_t = \alpha^{-1}(\widehat{y}_t - \widehat{\xi}_{a,t})$ into (n) to find $\widehat{w}_t = (\frac{\varphi}{\alpha} + \sigma)\widehat{y}_t - \widehat{\xi}_{c,t} - \frac{\varphi}{\alpha}\widehat{\xi}_{a,t}$. Inserting these expressions into the marginal cost term in the Phillips-curve yields this term as

$$\begin{aligned}(\tfrac{\varphi}{\alpha} + \sigma)\widehat{y}_t - \widehat{\xi}_{c,t} - (\tfrac{\varphi}{\alpha} + 1)\widehat{\xi}_{a,t} + \tfrac{1-\alpha}{\alpha}(\widehat{y}_t - \widehat{\xi}_{a,t}) &= \frac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widehat{y}_t - \widehat{\xi}_{c,t} - \frac{1+\varphi}{\alpha}\widehat{\xi}_{a,t} \\ &= \frac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widetilde{y}_t,\end{aligned}$$

so that we receive the Phillips-curve as

$$\pi_t = \beta\mathbb{E}_t\pi_{t+1} + \underbrace{\frac{\epsilon-1}{\kappa}\frac{1+\varphi+\alpha(\sigma-1)}{\alpha}}_{=:\ \theta}\widetilde{y}_t - \frac{1}{\kappa}\widehat{\xi}_{\tau,t}.$$

Finally, we can add and subtract $\widehat{y}_t, \widehat{y}_{t+1}$ in equation (b) to receive the canonical IS-equation. This then yields the canonical representation of the NK-model:

$$\boxed{\begin{array}{lc} (PC) & \pi_t = \beta\mathbb{E}_t\pi_{t+1} + \theta\widetilde{y}_t - \frac{1}{\kappa}\widehat{\xi}_{\tau,t} \\ (IS) & \widetilde{y}_t = \mathbb{E}_t\widetilde{y}_{t+1} - \frac{1}{\sigma}\left(i_t - \mathbb{E}_t\pi_{t+1} + \ln\beta - r_t^e\right) \end{array}}$$

where r_t^e is the real rate consistent with the efficient allocation, see above.

Remark 1 (Alternative derivation of the Allocation Gap law of motion)**.** *The above represents the 'textbook way' of deriving the allocation gap law of motion. For more complicated systems this can become quite tedious. A derivation that always works is first subtracting matching equations (i.e. subtract* $(n)^e$ *from* (n)*, and so on) and then simplify as far as possible from there on.*

Doing this for the above model (multiply $(PC)^e$ *by* $\frac{\epsilon-1}{\kappa}$ *first) produces the system*

$$
\begin{aligned}
&(n)-(n)^e \quad && \varphi\widetilde{n}_t+\sigma\widetilde{y}_t=\widetilde{w}_t\\
&(b)-(b)^e \quad && \iff (IS)\\
&(PC)-(PC)^e \quad && \pi_t=\beta\mathbb{E}_t\pi_{t+1}+\frac{\epsilon-1}{\kappa}(\widetilde{w}_t+(1-\alpha)\widetilde{n}_t)-\frac{1}{\kappa}\widehat{\xi}_{\tau,t}\\
&(y)-(y)^e \quad && \widetilde{y}_t=\alpha\widetilde{n}_t
\end{aligned}
$$

and from here on, straightforward substitutions produce the familiar two-equation system from above. ⋄

3 Loss Function

Since the non-stochastic steady state of the here-presented model is efficient, we may perform a naive LQ approximation to the Ramsey program. We have already derived the linear approximation to the implementability constraints, it only remains to derive the quadratic approximation to the Ramsey-planner's objective function which is the utility function of the representative agent.

Said utility function is additively separable across time, so we concentrate on the summand for some $t \geq 0$. We have:

$$
\begin{aligned}
u_t &= \xi_{c,t}\frac{c_t^{1-\sigma}}{1-\sigma}-\chi\frac{n_t^{1+\varphi}}{1+\varphi}+\xi_{h,t}\frac{h_t^{1-\nu}}{1-\nu}\\
&= \xi_{c,t}\frac{\left(y_t(1-\frac{\kappa}{2}\pi_t^2)\right)^{1-\sigma}}{1-\sigma}-\chi\frac{\left(\frac{y_t}{\xi_{a,t}}\right)^{\frac{1+\varphi}{\alpha}}}{1+\varphi}+\xi_{h,t}\frac{h_{ss}^{1-\nu}}{1-\nu} \quad \text{(using market-clearing conditions).}
\end{aligned}
$$

The relevant derivatives are:

$$
\begin{aligned}
\partial u_t/\partial \pi_t &= \xi_{c,t}\left(y_t(1-\tfrac{\kappa}{2}\pi_t^2)\right)^{-\sigma}(-\kappa\pi_t y_t) \\
\partial u_t/\partial \pi_t|_{ss} &= 0 \\
\partial u_t^2/\partial \pi_t^2 &= \xi_{c,t}(-\sigma)\left(y_t(1-\tfrac{\kappa}{2}\pi_t^2)\right)^{-\sigma-1}(-\kappa\pi_t y_t)^2 - \xi_{c,t}\left(y_t(1-\tfrac{\kappa}{2}\pi_t^2)\right)^{-\sigma}\kappa y_t \\
\partial u_t^2/\partial \pi_t^2|_{ss} &= -y_{ss}^{1-\sigma}\kappa \\
\partial u_t/\partial y_t &= \xi_{c,t}\left(y_t(1-\tfrac{\kappa}{2}\pi_t^2)\right)^{-\sigma}(1-\tfrac{\kappa}{2}\pi_t^2) - \frac{\chi}{\alpha}(\tfrac{y_t}{\xi_{a,t}})^{\frac{1+\varphi-\alpha}{\alpha}}\xi_{a,t}^{-1} \\
\partial u_t/\partial y_t|_{ss} &= y_{ss}^{-\sigma} - \frac{\chi}{\alpha}(y_{ss})^{\frac{1+\varphi-\alpha}{\alpha}} = 0 \quad \text{(use } (n) \text{ in SS to see this.)} \\
\partial u_t^2/\partial y_t^2 &= \xi_{c,t}(-\sigma)\left(y_t(1-\tfrac{\kappa}{2}\pi_t^2)\right)^{-\sigma-1}(1-\tfrac{\kappa}{2}\pi_t^2)^2 - \frac{\chi}{\alpha}\frac{1+\varphi-\alpha}{\alpha}(\tfrac{y_t}{\xi_{a,t}})^{\frac{1+\varphi-\alpha}{\alpha}-1}\xi_{a,t}^{-2} \\
\partial u_t^2/\partial y_t^2|_{ss} &= -\sigma y_{ss}^{-\sigma-1} - \frac{\chi}{\alpha}y_{ss}^{\frac{1+\varphi-\alpha}{\alpha}}\frac{1+\varphi-\alpha}{\alpha}y_{ss}^{-1} \\
&= -y_{ss}^{-\sigma-1}\frac{1+\varphi+\alpha(\sigma-1)}{\alpha} \\
\partial u_t^2/\partial y_t\partial\xi_{c,t}|_{ss} &= y_{ss}^{-\sigma} \\
\partial u_t^2/\partial y_t\partial\xi_{a,t} &= -\frac{\chi}{\alpha}\frac{1+\varphi-\alpha}{\alpha}(\tfrac{y_t}{\xi_{a,t}})^{\frac{1+\varphi-\alpha}{\alpha}-1}\tfrac{y_t}{\xi_{a,t}^2}\xi_{a,t}^{-1}(-1) - \frac{\chi}{\alpha}(\tfrac{y_t}{\xi_{a,t}})^{\frac{1+\varphi-\alpha}{\alpha}}(-1)\xi_{a,t}^{-2} \\
\partial u_t^2/\partial y_t\partial\xi_{a,t}|_{ss} &= y_{ss}^{-\sigma}\frac{1+\varphi}{\alpha}
\end{aligned}
$$

This establishes that

$$
\begin{aligned}
u_t - u_{ss} &= -y_{ss}^{1-\sigma}\frac{\kappa}{2}\pi_t^2 - y_{ss}^{-\sigma-1}\frac{1+\varphi+\alpha(\sigma-1)}{2\alpha}y_{ss}^2\widehat{y}_t^2 \\
&\quad + y_{ss}^{-\sigma}\widehat{\xi}_{c,t}y_{ss}\widehat{y}_t + y_{ss}^{-\sigma}\frac{1+\varphi}{\alpha}\widehat{\xi}_{a,t}y_{ss}\widehat{y}_t + O(3) + \text{t.i.p.s.} \\
\implies u_t - u_{ss} &\propto -\kappa\pi_t^2 - \frac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widehat{y}_t^2 + 2\widehat{\xi}_{c,t}\widehat{y}_t + 2\frac{1+\varphi}{\alpha}\widehat{\xi}_{a,t}\widehat{y}_t + O(3) + \text{t.i.p.s.}
\end{aligned}
$$

Now observe that

$$
\begin{aligned}
&-\tfrac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widehat{y}_t^2 + 2\widehat{\xi}_{c,t}\widehat{y}_t + 2\tfrac{1+\varphi}{\alpha}\widehat{\xi}_{a,t}\widehat{y}_t = -\tfrac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widehat{y}_t\cdot\left(\widehat{y}_t - 2\left(\tfrac{\alpha}{1+\varphi+\alpha(\sigma-1)}\widehat{\xi}_{c,t} + \tfrac{1+\varphi}{1+\varphi+\alpha(\sigma-1)}\widehat{\xi}_{a,t}\right)\right) \\
&= -\tfrac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widehat{y}_t\cdot(\widehat{y}_t - 2\widehat{y}_t^e) \\
&= -\tfrac{1+\varphi+\alpha(\sigma-1)}{\alpha}\left[\widehat{y}_t\cdot(\widehat{y}_t - 2\widehat{y}_t^e) + (\widehat{y}_t^e)^2 - (\widehat{y}_t^e)^2\right] \\
&= -\tfrac{1+\varphi+\alpha(\sigma-1)}{\alpha}\widetilde{y}_t^2 + \text{t.i.p.s.},
\end{aligned}
$$

So that we have the loss function as

$$
\boxed{-(\mathcal{U} - \mathcal{U}|_{ss}) \propto \mathcal{L} + O(3) + \text{t.i.p.s.},\ \ \mathcal{L} := \mathbb{E}_0\sum_{t\geq 0}\beta^t\left\{\pi_t^2 + \tfrac{\theta}{\epsilon-1}\widetilde{y}_t^2\right\}}
$$

4 LQ-approximated Ramsey Program

The monetary authority thus solves

$$\min_{(\pi_t, \widetilde{y}_t, i_t)_{t\geq 0}} \mathcal{L} \quad \text{subject to}$$

$$(PC) \qquad \pi_t = \beta \mathbb{E}_t \pi_{t+1} + \theta \widetilde{y}_t - \tfrac{1}{\kappa} \widehat{\xi}_{\tau,t},$$

$$(IS) \qquad \widetilde{y}_t = \mathbb{E}_t \widetilde{y}_{t+1} - \tfrac{1}{\sigma} \left(i_t - \mathbb{E}_t \pi_{t+1} + \ln \beta - r_t^e \right).$$

Now since we ignore the ZLB-constraint on i_t, it is always possible to select i_t such that (IS) holds. Thus, we can drop i_t as a control and (IS) as a constraint. Then, we write the program in Lagrangian form (strong duality holds since this is a convex program) as

$$\min_{(\pi_t, \widetilde{y}_t)_{t\geq 0}} \max_{(\mu_t)_{t\geq 0}} \mathbb{E}_0 \sum_{t\geq 0} \beta^t \left\{ \pi_t^2 + \tfrac{\theta}{\epsilon - 1} \widetilde{y}_t^2 + 2\mu_t \cdot \left(\pi_t - \beta \pi_{t+1} - \theta \widetilde{y}_t + \tfrac{1}{\kappa} \widehat{\xi}_{\tau,t} \right) \right\},$$

where we have directly omitted the $\mathbb{E}_t$-operator inside the summation because of the law of iterated expectations. The first-order conditions are $\forall t \geq 0$ and with the pre-commitment μ_{-1} given:

$$\boxed{(\pi)\ \pi_t + \mu_t - \mu_{t-1} = 0, \quad (\widetilde{y})\ \widetilde{y}_t = (\epsilon - 1)\mu_t, \quad (\mu)\ \pi_t = \beta \mathbb{E}_t \pi_{t+1} + \theta \widetilde{y}_t - \tfrac{1}{\kappa} \widehat{\xi}_{\tau,t}.}$$

We may derive two canonical results:

1. **Targeting Rule**: Optimal monetary policy maintains a constant and negative relationship between the growth rate of the output gap and that of inflation:

$$\pi_t + \frac{\Delta \widetilde{y}_t}{\epsilon - 1} = 0.$$

In doing so, optimal monetary policy will keep the output gap on a downward trajectory as long as inflation is above its unconstrained-optimal level of zero. This makes sense when looking at (PC): as long as inflation is "too high", we can decrease it by shifting down the output gap today, or the output gap tomorrow (just iterate (PC) forward). Symmetrically, OMP will implement an upward trajectory of the output gap if inflation is below target. At its core, the dynamics that this equation encompasses illustrates how commitment works. Take the targeting rule in $t+1$: $\pi_{t+1} + \mu_{t+1} - \mu_t = 0$. Ceteris paribus, lowering inflation in $t+1$ will, besides affecting $\mathcal{L}$ with slope π_{t+1}, downward-violate the $t+1$ Phillips curve, $\pi_{t+1} - \beta \pi_{t+2} < 0$ c.p.; this gets punished at the margin with rate μ_{t+1}; but it also *upward*-violates the time t Phillips curve, $\pi_t - \beta \pi_{t+1} > 0$ c.p.; and this means that inflation in t may be lowered, which may be beneficial depending on where we start in terms of π_t. The overall effect of these dynamic linkages will become clearer once we have solved for the process $(\mu_t)_t$ that solves the FOC.

2. **Divine Coincidence**: In the absence of cost-push shocks, $\widehat{\xi}_\tau \equiv 0$, optimal monetary policy may achieve the first-best allocation, $\pi_t = \widetilde{y}_t = 0$, $\forall t$ almost surely, by "tracking the efficient rate", i.e. by setting $i_t + \ln\beta = r^e_t$. This may, e.g., be implemented by the rule $i_t = -\ln\beta + r^e_t + \phi_\pi \pi_t$ where $\phi_\pi > 1$ ensures that the first best allocation is the unique market equilibrium (this is the Taylor Principle).

The FOC also admit a unique explicit solution. Before characterizing this solution, we prove that there exists indeed a unique non-explosive solution. The requirement of non-explosivity is a consequence of an optimality condition that we did not explicitly state: the transversality condition.

As a first step, we obtain the law of motion of μ_t by inserting (π) and ($\widehat{y}$) into (μ), thus obtaining

$$\mu_{t-1} - \mu_t = \beta\mathbb{E}_t\{\mu_t - \mu_{t+1}\} + \theta(\epsilon - 1)\mu_t - \tfrac{1}{\kappa}\widehat{\xi}_{\tau,t}.$$

Now defining $\boldsymbol{\mu}_t := (\mu_t, \mu_{t-1})^\top$ allows to write the above law of motion as a vector-valued first-order expectational difference equation:

$$\beta\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}\mathbb{E}_t\boldsymbol{\mu}_{t+1} = \begin{pmatrix}1+\theta(\epsilon-1) & -1\\ \beta & 0\end{pmatrix}\boldsymbol{\mu}_t - \begin{pmatrix}\frac{1}{\kappa}\widehat{\xi}_{\tau,t}\\ 0\end{pmatrix}$$

which is equivalent to

$$\mathbb{E}_t\boldsymbol{\mu}_{t+1} = \beta^{-1}\underbrace{\begin{pmatrix}1+\theta(\epsilon-1)+\beta & -1\\ \beta & 0\end{pmatrix}}_{=:\,\underline{\boldsymbol{A}}}\boldsymbol{\mu}_t - \begin{pmatrix}\frac{1}{\beta\kappa}\widehat{\xi}_{\tau,t}\\ 0\end{pmatrix}.$$

The Eigenvalues $e_{+,-}$ of $\underline{\boldsymbol{A}}$ solve $e^2 - e\,\mathrm{tr}\,\underline{\boldsymbol{A}} + \det\underline{\boldsymbol{A}} = 0 \iff e^2 - e\cdot(1+\theta(\epsilon-1)+\beta)+\beta = 0$. The larger Eigenvalue is

$$e_+ = \tfrac{1}{2}\left[1+\theta(\epsilon-1)+\beta+\sqrt{(1+\theta(\epsilon-1)+\beta)^2-4\beta}\right] > \tfrac{1}{2}\left[2(1+\theta(\epsilon-1)+\beta)-2\sqrt{\beta}\right] = 1+\theta(\epsilon-1)+\beta-\sqrt{\beta}.$$[1]

Thus, the largest Eigenvalue of $\beta^{-1}\underline{\boldsymbol{A}}$ is larger than $\beta(1-\sqrt{\beta}+\theta(\epsilon-1)+\beta) > \beta^{-1}\beta = 1$. Therefore, the transition matrix of the law of motion of $\boldsymbol{\mu}$ has an explosive Eigenvalue which means that any solution that deviates from the canonical particular solution must explode. Therefore, the stable solution is unique.

To find the stable solution, we guess that it is of an $AR(1)$ form and apply the method of undetermined coefficients. In doing so, we assume w.l.o.g. that $\xi_{\tau,t}$ is i.i.d.[2] That is, we

[1] The first inequality is an application of: $\forall x > y > 0,\ \sqrt{x-y} > \sqrt{x} - \sqrt{y}$.

[2] This is w.l.o.g. because the dynamics of μ that we solve for capture the MA-coefficients of the general solution. If we want to construct a solution for μ for a general covariance-stationary process ξ_τ, the appropriate setup for the method of undetermined coefficients is to guess that the law of motion for μ takes the form $\mu_t = a\mu_{t-1} + \frac{1}{\kappa}\sum_{s\geq 0} b_s\mathbb{E}_t\widehat{\xi}_{\tau,t+s}$, with a, $(b_s)_{s=0}^\infty$ being the coefficients to be determined. Plugging into $\mu_{t-1}+\zeta\mu_t+\beta\mathbb{E}_t\mu_{t+1}+\frac{1}{\kappa}\xi_{\tau,t} = 0$ and isolating terms delivers the restrictions $1+\zeta a+\beta a^2 = 0$, $b_0 = -(\zeta+a\beta)^{-1}$, $\forall s \geq 1$, $b_s = -\beta(\zeta+a\beta)^{-1}b_{s-1}$. So, the coefficients a, b_0 are the same as when ξ_τ is i.i.d. (in particular, both are between 0 and 1), and the remaining coefficients are geometrically

guess that the solution is of the form

$$\mu_t = a\mu_{t-1} + \tfrac{b}{\kappa}\widehat{\xi}_{\tau,t}$$

with a, b coefficients to be determined. We substitute this solution into the law of motion for μ_t until we have only time $t-1$ or exogenous variables left; defining $\zeta := -(1+\beta+\theta(\epsilon-1))$ we get:

$$\begin{aligned} \mu_{t-1} + \zeta\mu_t + \beta\mathbb{E}_t\mu_{t+1} + \tfrac{1}{\kappa}\xi_{\tau,t} &= 0 \\ \iff \mu_{t-1}\cdot[1+\zeta a+\beta a^2] + \tfrac{1}{\kappa}\xi_{\tau,t}\cdot[1+\zeta b+\beta ab] &= 0 \\ \Longleftarrow a \in \left\{\frac{-\zeta \pm \sqrt{\zeta^2-4\beta}}{2\beta}\right\}, \quad b = -(\zeta+\beta a)^{-1}. \end{aligned}$$

Now since $\frac{-\zeta+\sqrt{\zeta^2-4\beta}}{2\beta} > 1$, as seen above, it must be that the AR-coefficient we are looking for is

$$0 < a = \frac{-\zeta - \sqrt{\zeta^2-4\beta}}{2\beta} < \frac{1+\beta+\theta(\epsilon-1)-(1-\beta+\theta(\epsilon-1))}{2\beta} = 1.^3$$

Now since $a \in (0,1)$ we also have $b = (1+\beta(1-a)+\theta(\epsilon-1))^{-1} \in (0,1)$. This finally leads to

Theorem 1 (Optimal Monetary Policy)**.** *Given the stated environment, Ramsey-optimal monetary policy under commitment is characterized by the following features*

1. *In the absence of a cost-push shock, OMP tracks the efficient rate by setting*

$$i_t = -\ln\beta + r_t^e$$

where

$$r_t^e := \sigma\mathbb{E}_t\Delta\widehat{y}_{t+1}^e - \mathbb{E}_t\Delta\widehat{\xi}_{c,t+1}.$$

That is, OMP (i) increases the nominal rate if the efficient output is expected to rise (to prevent actual output today from rising as well, because of income-smoothing effects) and OMP (ii) c.p. increases the nominal rate if the consumption taste process is expected to fall (to prevent actual output from increasing today, because of preference-driven intertemporal substitution; this is on top of the change implied through the first difference of efficient output.)[4]

decreasing as $s \to \infty$. A more explicit characterization requires detailing $\mathbb{E}_t\xi_{\tau,t+s}$ by specifying a law of motion for ξ_τ.

[3]Details: $a > 0$ is obvious; furthermore, it is $\zeta^2 - 4\beta = 1+2\beta+\beta^2 + 2\theta(\epsilon-1)+2\beta\theta(\epsilon-1) + \theta^2(\epsilon-1)^2 - 4\beta = 1-2\beta+\beta^2 + 2\theta(\epsilon-1)(1-\beta)+4\beta\theta(\epsilon-1) + \theta^2(\epsilon-1)^2 = [1-\beta+\theta(\epsilon-1)]^2+4\beta\theta(\epsilon-1)$. Now since $\theta > 0$, $\epsilon > 1$, we have $\zeta^2-4\beta > [1-\beta+\theta(\epsilon-1)]^2$ which, by monotonicity of $x \mapsto \sqrt{x}$, proves the second inequality.

[4]This point is quite nuanced: one would expect OMP to accommodate temporary changes in preferences,

2. *In response to a one-off cost-push shock, $\widehat{\xi}_{\tau,0} < 0, (\widehat{\xi}_{\tau,t})_{t\geq 1} = 0$, OMP allows the on-impact output gap to be slightly negative, i.e. the elasticity of the output gap on impact with respect to the shock is $b/\kappa \in (0, 1/\kappa)$. Since the output gap drops on impact, inflation rises on impact by $\frac{b}{\kappa(\epsilon-1)}$. With passing time after the shock, $t \geq 0$, OMP geometrically increases the output gap back to zero, with coefficient $a \in (0, 1)$. This means that the first differences of the output gap are positive and decreasing, meaning that inflation in $t \geq 1$ is negative and mean-reverting. Thus, optimal monetary policy under commitment achieves a milder inflation-output trade-off than under discretion by promising to keep the output gap negative for some time, thus decreasing inflation today not only by a negative output gap but also by a negative expected inflation rate.*

as here ξ_c. The fact that such increases in ξ_c directly affect the efficient rate and that OMP offsets this intertemporal substitution pressure towards consumption tomorrow with an interest rate hike is reflective of the fact that such an intertemporal substitution is *impossible* in the given economy. Since there is no means to shift resources intertemporally, the optimal monetary policy undoes the desire to perform such a shift. The only way ξ_c can actually impact the allocation obtained under OMP is by changing the efficient level of output through its impact on the marginal rate of substitution between leisure and consumption. An increase in ξ_c acts like a negative income effect on labor supply (by increasing the marginal utility of consumption). And since shifting resources *intra*temporally between labor and consumption *is* possible, OMP will reflect the ensuing shift in the marginal rate of substitution between consumption and leisure in its chosen allocation and optimally increase the level of output by decreasing the nominal rate. (A mean-reverting increase in ξ_c will produce a *negative* expected growth rate of efficient output.)