Topics in Economic Policy
Supervision 3

Samuel Lee

Question 1

To reiterate, the output gap is

$$y = \alpha - (\kappa + \varepsilon)r + \upsilon$$

where r is the policy instrument, α is an anticipated shock known by the policymaker when setting r, κ is a positive parameter that I assume is known by the policymaker, and ε and υ are independent shocks not observed by the policymaker with variances σ_ε^2 and σ_υ^2 (I assume σ_κ^2 is a typo).

The policymaker faces both multiplicative uncertainty from ε and additive uncertainty from υ when minimising $L = \mathrm{E}[y^2]$. The first-order condition for a minimum is

$$\begin{aligned}
\frac{\partial L}{\partial r} &= -2\,\mathrm{E}[y(\kappa + \varepsilon)] = 0 \\
\mathrm{E}[\alpha(\kappa + \varepsilon) - (\kappa + \varepsilon)^2 r + \upsilon(\kappa + \varepsilon)] &= 0 \\
\alpha\kappa - (\kappa^2 + \sigma_\varepsilon^2)r &= 0 \\
r &= \frac{\alpha\kappa}{\kappa^2 + \sigma_\varepsilon^2}
\end{aligned}$$

which follows the standard result: multiplicative uncertainty ($\sigma_\varepsilon^2 > 0$) results in a conservative choice of r (the Brainard principle) whereas additive uncertainty ($\sigma_\upsilon^2 > 0$) does not affect the coefficients of the optimal policy instrument function (certainty equivalence).

Question 2

(a)

The policymaker observes all the shocks, and minimises

$$L_t = \lambda y_t^2 + \pi_t^2$$

Again, the first-order condition for a minimum is

$$\frac{\partial L_t}{\partial i_t} = 2\lambda y_t \frac{\partial y_t}{\partial i_t} + 2\pi_t \alpha \frac{\partial y_t}{\partial i_t} = -2\beta\lambda y_t - 2\alpha\beta\pi_t = 0$$

which implies

$$\begin{aligned}
\lambda y_t + \alpha\pi_t &= -\beta\lambda(i_t - \pi_t^e) + \lambda\eta_t + \alpha\pi_t^e + \alpha^2 y_t + \alpha\varepsilon_t = 0 \\
\beta\lambda(i_t - \pi_t^e) - \lambda\eta_t - \alpha\pi_t^e + \alpha^2\beta(i_t - \pi_t^e) - \alpha^2\eta_t - \alpha\varepsilon_t &= 0 \\
(\alpha^2+\lambda)\beta i_t - \left[\alpha + (\alpha^2+\lambda)\beta\right]\pi_t^e - (\alpha^2+\lambda)\eta_t - \alpha\varepsilon_t &= 0 \\
i_t &= \left[1 + \frac{\alpha}{(\alpha^2+\lambda)\beta}\right]\pi_t^e + \frac{1}{\beta}\eta_t + \frac{\alpha}{(\alpha^2+\lambda)\beta}\varepsilon_t
\end{aligned}$$

which shows that the nominal interest rate rises more than one-for-one with expected inflation, aligning with the Taylor principle. Taking expectations,

$$\begin{aligned}
\mathrm{E}[i_t] &= \left[1 + \frac{\alpha}{(\alpha^2+\lambda)\beta}\right]\pi_t^e \\
\mathrm{E}[y_t] &= -\beta(\mathrm{E}[i_t] - \pi_t^e) \\
\mathrm{E}[\pi_t] &= \pi_t^e + \alpha\,\mathrm{E}[y_t]
\end{aligned}$$

The last equation implies $\mathrm{E}[y_t] = 0$. $\mathrm{E}[y_t] = 0$ implies that $\mathrm{E}[i_t] = \pi_t^e$. $\mathrm{E}[i_t] = \pi_t^e$ implies $\pi_t^e = 0$. So we have the optimal interest rate rule

$$i_t = \underbrace{\frac{1}{\beta}}_{\theta}\eta_t + \underbrace{\frac{\alpha}{(\alpha^2+\lambda)\beta}}_{\gamma}\varepsilon_t$$

which could mean that the nominal interest rate is set higher in response to positive aggregate demand shocks (η_t) and cost-push shocks (ε_t).

The higher β is, the smaller the increase in i_t required to effect the same reduction in y_t. This means the interest rate response to a positive η_t can now be smaller, but it also means that the interest rate response to a positive ε_t becomes smaller as controlling inflation becomes more costly.

The responsiveness of interest rates to ε changes with α, but the direction is ambiguous:

$$\frac{\partial}{\partial\alpha}\frac{\alpha}{(\alpha^2+\lambda)\beta} = \frac{1}{(\alpha^2+\lambda)\beta} - \frac{2\alpha^2\beta}{[(\alpha^2+\lambda)\beta]^2} = \frac{1}{(\alpha^2+\lambda)\beta} - \frac{2\alpha^2}{(\alpha^2+\lambda)^2\beta} = \frac{\lambda-\alpha^2}{(\alpha^2+\lambda)^2\beta}$$

which means the interest rates are most responsive to ε when $\alpha = \sqrt{\lambda}$. Increases in α essentially do two things:

1. Make reductions in inflation in response to a positive ε_t less costly in terms of y_t
2. Reduce the interest rate hike required to bring inflation down by a given level

The first effect makes the policymaker more willing to raise interest rates, while the second effect allows the policymaker to get away with raising interest rates by a smaller amount. Evidently, at lower levels of α, effect 1 dominates and at higher levels of α, effect 2 dominates.

(b)

We can now derive the ex-post y_t and π_t:

$$y_t = -\beta\left[\frac{1}{\beta}\eta_t + \frac{\alpha}{(\alpha^2+\lambda)\beta}\varepsilon_t\right] + \eta_t = -\frac{\alpha}{\alpha^2+\lambda}\varepsilon_t$$
$$\pi_t = \alpha y_t + \varepsilon_t = \frac{\lambda}{\alpha^2+\lambda}\varepsilon_t$$

using $\pi_t^e = 0$ and the solved value of i_t. With that, the volatility of the outcome variables are

$$\sigma_y^2 = \mathrm{Var}[y_t] = \frac{\alpha^2}{(\alpha^2+\lambda)^2}\sigma_\varepsilon^2 \implies \sigma_y = \sqrt{\mathrm{Var}[y_t]} = \frac{\alpha}{\alpha^2+\lambda}\sigma_\varepsilon$$
$$\sigma_\pi^2 = \mathrm{Var}[\pi_t] = \frac{\lambda^2}{(\alpha^2+\lambda)^2}\sigma_\varepsilon^2 \implies \sigma_\pi = \sqrt{\mathrm{Var}[\pi_t]} = \frac{\lambda}{\alpha^2+\lambda}\sigma_\varepsilon$$

which means that

$$\sigma_\pi^2 = \left(\sqrt{\sigma_\varepsilon^2} - \alpha\sqrt{\sigma_y^2}\right)^2$$

since

$$\left(\sqrt{\sigma_\varepsilon^2} - \alpha\sqrt{\sigma_y^2}\right)^2 = \left(\frac{\alpha^2+\lambda}{\alpha^2+\lambda}\sigma_\varepsilon - \alpha\frac{\alpha}{\alpha^2+\lambda}\sigma_\varepsilon\right)^2 = \frac{\lambda^2}{\alpha^2+\lambda}\sigma_\varepsilon^2 = \sigma_\pi^2$$

We can see that η_t does not feature in any of the final outcomes, whereas ε_t seems to result in a trade-off for the policymaker. The reason is that η_t doesn't affect the choice set of the policymaker since they observe the shock and can offset it exactly by adding $\frac{1}{\beta}\eta_t$ to whatever interest rate they would have set if $\eta_t = 0$; they have complete control over what y_t they wish to see. The problem with ε_t is that it features in the Phillips relation, which governs what combinations of (π_t, y_t) are feasible.

(c)

If the volatility is due to greater σ_η^2 and σ_ε^2, it doesn't affect the coefficients of the policy rule, since the shocks are additive and don't change the marginal effect of i_t on the loss function (or we can just see that the derived policy rule ignores volatility). But inflation and output could become more volatile for other reasons: model uncertainty or uncertainty in model parameters like α and β. Those would affect the policy rule since they affect the expected loss function in a non-linear way.

Question 3

(a)

The equations are exactly as in Question 2, with $r_t = i_t - \pi_t^e$. (1) is an goods market equilibrium condition, (2) is a Phillips relation, and (3) is a loss function with a bliss point at zero inflation and no output gap.

(b)

The derivation will be identical to Question 2 except that $r_t = i_t - \pi_t^e$. We found that

$$i_t = \left[1 + \frac{\alpha}{(\alpha^2 + \lambda)\beta}\right]\pi_t^e + \frac{1}{\beta}\eta_t + \frac{\alpha}{(\alpha^2 + \lambda)\beta}\varepsilon_t$$

So we get r_t by subtracting π_t^e on both sides, and noting that with rational expectations $\pi_t^e = 0$ we get

$$r_t = \frac{1}{\beta}\eta_t + \frac{\alpha}{(\alpha^2 + \lambda)\beta}\varepsilon_t$$

and any comments would be the same as before.

(c)

All shocks are independent, and η_t, $\eta_{S,t}$, and $\eta_{IMF,t}$ are jointly normal. We might guess that the conditional expectation of η_t given $\eta_{S,t}$ and $\eta_{IMF,t}$ is linear (this is indeed true, as shown in the endnote, although deriving this makes the following redundant). Since the conditional expectation of some random variable is the best forecast in a mean squared error sense, if we are right that the conditional expectation is linear, we can find the optimal forecast by finding the best linear predictor of η_t. Using the law of iterated expectations, we know that the expectation of the optimal forecast must be the expectation of η_t, so

$$\mathrm{E}[\phi + \phi_S \eta_{S,t} + \phi_{IMF}\eta_{IMF,t}] = \phi + (\phi_S + \phi_{IMF})\,\mathrm{E}[\eta_t] = \mathrm{E}[\eta_t]$$

which implies $\phi = (1 - \phi_S - \phi_{IMF})\,\mathrm{E}[\eta_t] = 0$. *Despite how the lecture notes work through a similar problem, I believe this in no way implies any restriction on the sum of ϕ_S and ϕ_{IMF}.*

We can then minimise the mean squared error over all the linear predictors. Any linear predictor is also unbiased as $\mathrm{E}[\eta_t] = \mathrm{E}[v_{S,t}] = \mathrm{E}[v_{IMF,t}] = 0$, so the mean squared error is just the variance:

$$\begin{aligned}\mathrm{E}[(\phi_S \eta_{S,t} + \phi_{IMF}\eta_{IMF,t} - \eta_t)^2] &= \mathrm{Var}[\phi_S v_{S,t} + \phi_{IMF} v_{IMF,t} - (1 - \phi_S - \phi_{IMF})\eta_t] \\ &= \phi_S^2\sigma_S^2 + \phi_{IMF}^2\sigma_{IMF}^2 + (1 - \phi_S - \phi_{IMF})^2\sigma_\eta^2\end{aligned}$$

The first-order conditions for a minimum are

$$\left.\begin{aligned}\phi_S\sigma_S^2 &= (1 - \phi_S - \phi_{IMF})\sigma_\eta^2 \\ \phi_{IMF}\sigma_{IMF}^2 &= (1 - \phi_S - \phi_{IMF})\sigma_\eta^2\end{aligned}\right\} \implies \phi_{IMF} = \frac{\sigma_S^2}{\sigma_{IMF}^2}\phi_S$$

which implies

$$\phi_S\sigma_S^2 = \left(1 - \phi_S - \frac{\sigma_S^2}{\sigma_{IMF}^2}\phi_S\right)\sigma_\eta^2 \implies \phi_S = \frac{\sigma_\eta^2}{\sigma_S^2 + \sigma_\eta^2 + \frac{\sigma_S^2\sigma_\eta^2}{\sigma_{IMF}^2}} = \frac{\sigma_{IMF}^2}{\frac{\sigma_S^2\sigma_{IMF}^2}{\sigma_\eta^2} + \sigma_S^2 + \sigma_{IMF}^2}$$

$$\phi_{IMF} = \frac{\sigma_S^2}{\frac{\sigma_S^2\sigma_{IMF}^2}{\sigma_\eta^2} + \sigma_S^2 + \sigma_{IMF}^2}$$

The more noisy η_{IMF} is, the greater the weight we place on η_S. The more noisy η_S is, the greater the weight we place on η_{IMF}. And the greater the variance of η, the more we will want to rely on the signals from η_S and η_{IMF}, but when the distribution of η is very concentrated with a low variance, we do better by giving a forecast close to 0 since that's where the true value is likely to be and we wouldn't put much weight on values of υ_S and υ_{IMF} that may possibly stray very far.

(d)

We now minimise $\mathrm{E}[L_t]$ conditional on the policymaker's information $\Omega_P = \{\pi_t^e, \varepsilon_t, \eta_{S,t}, \eta_{IMF,t}, \Omega\}$ where $\Omega = \{\alpha, \beta\}$. We have

$$\mathrm{E}[L_t|\Omega_P] = \mathrm{E}[\pi_t^2 + \lambda y_t^2|\Omega_P]$$

Differentiating through the expectation, we get

$$\begin{aligned}
\frac{\partial}{\partial r_t}\mathrm{E}[L_t|\Omega_P] &= \mathrm{E}\left[2\pi_t\frac{\partial \pi_t}{\partial y_t}\frac{\partial y_t}{\partial r_t} + 2\lambda y_t \frac{\partial y_t}{\partial r_t} \;\middle|\; \Omega_P\right] \\
&= \mathrm{E}[-2\alpha\beta(\pi_t^e + \alpha y_t + \varepsilon_t) - 2\beta\lambda(-\beta r_t + \eta_t)|\Omega_P] \\
&= -2\beta\,\mathrm{E}[\alpha\pi_t^e + \alpha^2(-\beta r_t + \eta_t) + \alpha\varepsilon_t - \beta\lambda r_t + \lambda\eta_t|\Omega_P]
\end{aligned}$$

Setting the above to zero, we get

$$\begin{aligned}
\mathrm{E}[\alpha\pi_t^e - (\alpha^2+\lambda)\beta r_t + (\alpha^2+\lambda)\eta_t + \alpha\varepsilon_t|\Omega_P] &= 0 \\
\alpha\pi_t^e - (\alpha^2+\lambda)\beta r_t + (\alpha^2+\lambda)\,\mathrm{E}[\eta_t|\Omega_P] + \alpha\varepsilon_t &= 0 \\
r_t = \frac{\alpha}{(\alpha^2+\lambda)\beta}\pi_t^e + \frac{1}{\beta}\mathrm{E}[\eta_t|\Omega_P] &+ \frac{\alpha}{(\alpha^2+\lambda)\beta}\varepsilon_t
\end{aligned}$$

and by now $\pi_t^e = 0$ should not be controversial, so the optimal rate is

$$r_t = \frac{1}{\beta}\mathrm{E}[\eta_t|\Omega_P] + \frac{\alpha}{(\alpha^2+\lambda)\beta}\varepsilon_t$$

which exhibits certainty equivalence: the optimal rate is the same as the rate derived under no uncertainty, except that η is replaced with its conditional expectation. This is the standard result for additive uncertainty. If the conditional expectation is linear, then this is simply the variance-weighted sum of $\eta_{S,t}$ and $\eta_{IMF,t}$ with the weights we derived earlier.

Note (Question 3c)

We can derive $\mathrm{E}[\eta_t|\eta_{S,t}, \eta_{IMF,t}]$. Suppressing time indices, we have that

$$\mathrm{E}[\eta|\eta_S, \eta_{IMF}] = \int_{-\infty}^{\infty} \eta \cdot f(\eta|\eta_S, \eta_{IMF})\,\mathrm{d}\eta$$

where we abuse notation to denote by $f(\eta|\eta_S, \eta_{IMF})$ the conditional distribution of η given values of η_S and η_{IMF}. We have

$$f(\eta|\eta_S, \eta_{IMF}) = \frac{f(\eta, \eta_S, \eta_{IMF})}{f(\eta_S, \eta_{IMF})} = \frac{f(\eta_S, \eta_{IMF}|\eta) f(\eta)}{f(\eta_S, \eta_{IMF})} = \frac{f(\eta_S, \eta_{IMF}|\eta) f(\eta)}{\int_{-\infty}^{\infty} f(\eta_S, \eta_{IMF}|\eta) f(\eta)\,\mathrm{d}\eta}$$

We know that η is normally distributed with zero mean and variance σ_η^2, and that given η, the remaining variations in η_S and η_{IMF} are explained by υ_S and υ_{IMF}, which are independent of one another and also have zero mean and variances σ_S^2 and σ_{IMF}^2. So $\eta_S \perp \eta_{IMF}|\eta$, which implies $f(\eta_S, \eta_{IMF}|\eta) = f(\eta_S|\eta)f(\eta_{IMF}|\eta)$, and we have in the numerator

$$f(\eta_S, \eta_{IMF}|\eta)f(\eta) = \frac{1}{\sqrt{2\pi\sigma_\eta^2}}e^{-\frac{\eta^2}{2\sigma_\eta^2}} \times \frac{1}{\sqrt{2\pi\sigma_S^2}}e^{-\frac{(\eta_S-\eta)^2}{2\sigma_S^2}} \times \frac{1}{\sqrt{2\pi\sigma_{IMF}^2}}e^{-\frac{(\eta_{IMF}-\eta)^2}{2\sigma_{IMF}^2}}$$

since for $X \in \{S, IMF\}$, the event $\eta_X = x$ given η has the same probability density as the event $\upsilon_X = x - \eta$.

Ignoring everything but the exponential terms, we have

$$\begin{aligned}
&\exp\left[-\frac{\sigma_S^2\sigma_{IMF}^2\eta^2 + \sigma_\eta^2\sigma_{IMF}^2(\eta_S-\eta)^2 + \sigma_\eta^2\sigma_S^2(\eta_{IMF}-\eta)^2}{2\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}\right] \\
&= \exp\left[-\frac{(\sigma_S^2\sigma_{IMF}^2 + \sigma_\eta^2\sigma_{IMF}^2 + \sigma_\eta^2\sigma_S^2)\eta^2 - 2\sigma_\eta^2(\sigma_{IMF}^2\eta_S + \sigma_S^2\eta_{IMF})\eta + \sigma_\eta^2(\sigma_{IMF}^2\eta_S^2 + \sigma_S^2\eta_{IMF}^2)}{2\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}\right] \\
&= \exp\left[-\frac{\eta^2 - 2\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\eta + \sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}}{2\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}}\right] \\
&= \exp\left[-\frac{\eta^2 - 2\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\eta + \left(\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\right)^2 - \left(\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\right)^2 + \sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}}{2\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}}\right] \\
&= \exp\left[-\frac{\left(\eta - \sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\right)^2}{2\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}}\right]\exp\left[\frac{\left(\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\right)^2 - \sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}}{2\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}}\right]
\end{aligned}$$

The first exponential term is proportional to the density function of a normally distributed variable with mean $\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}$ and variance $\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}$. The denominator of $f(\eta|\eta_S, \eta_{IMF})$ is simply the numerator integrated with respect to η over $(-\infty, \infty)$, so the second exponential term cancels out, the non-exponential terms we ignored before cancel out, and we are left with

$$f(\eta|\eta_S, \eta_{IMF}) = \frac{1}{\sqrt{2\pi\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}}}\exp\left[-\frac{\left(\eta - \sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}\right)^2}{2\frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}}\right]$$

which is now exactly the density function for a $N\left(\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S+\sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_{IMF}^2+\sigma_\eta^2\sigma_S^2}, \frac{\sigma_\eta^2\sigma_S^2\sigma_{IMF}^2}{\sigma_{IMF}^2\eta_S^2+\sigma_S^2\eta_{IMF}^2}\right)$ variable. The conditional expectation is therefore just the mean of that variable, which can also be expressed as

$$\sigma_\eta^2\frac{\sigma_{IMF}^2\eta_S + \sigma_S^2\eta_{IMF}}{\sigma_S^2\sigma_{IMF}^2 + \sigma_\eta^2\sigma_{IMF}^2 + \sigma_\eta^2\sigma_S^2} = \frac{\sigma_{IMF}^2}{\frac{\sigma_S^2\sigma_{IMF}^2}{\sigma_\eta^2} + \sigma_{IMF}^2 + \sigma_S^2}\eta_S + \frac{\sigma_S^2}{\frac{\sigma_S^2\sigma_{IMF}^2}{\sigma_\eta^2} + \sigma_{IMF}^2 + \sigma_S^2}\eta_{IMF}$$

which has an intuitive interpretation. The more noisy η_{IMF} is, the greater the weight we place on η_S. The more noisy η_S is, the greater the weight we place on η_{IMF}. And the greater the variance of η, the more we will want to rely on the signals from η_S and η_{IMF}, but when the distribution of η is very concentrated with a low variance, we do better by giving a forecast close to 0 since that's where the true value is likely to be and we wouldn't put much weight on values of η_S and η_{IMF} that may possibly stray very far.