Appendix

A. Approximations

We first demonstrate that EHS’ Taylor approximations may lead to wrong conclusions. Then, we present a better approximation, which does not lead to the wrong conclusions (i.e., one that delivers the same conclusions as obtained without approximations).\(^1\) Finally, we demonstrate why this approximation is better.

A.1. Demonstration that EHS’ Taylor approximations may lead to wrong conclusions

In order to demonstrate that EHS’ Taylor approximations are too crude, and therefore may lead to wrong conclusions, it suffices to consider the case of \(z = 0\). Hence, \(\pi^e = 0\) and (5) becomes:

\[
\pi = -\frac{b}{a_t + b^2 \varepsilon}. \tag{A.1}
\]

From this and (1), output follows as

\[
y = y^* + \frac{a_t}{a_t + b^2 \varepsilon}. \tag{A.2}
\]

\(^1\) We are very grateful to an anonymous referee who suggested this approximation.
In order to obtain expressions for the variance of inflation and output, EHS employ second-order approximations of ratios of stochastic variables around their respective means. Generally, they consider two independent stochastic variables $X$ and $Y$ with $\mu_X \equiv E[X]$ and $\mu_Y \equiv E[Y]$. A second-order Taylor approximation of $X/Y$ around $X = \mu_X$ and $Y = \mu_Y$ yields:

$$\frac{X}{Y} \approx \frac{X}{\mu_Y} - \frac{\mu_X}{(\mu_Y)^2} (Y - \mu_Y) + \frac{\mu_X}{(\mu_Y)^3} (Y - \mu_Y)^2 - \frac{1}{(\mu_Y)^2} (Y - \mu_Y) (X - \mu_X). \quad (A.3)$$

Hence, using the independence of $X$ and $Y$,

$$E \left[ \frac{X}{Y} \right] \approx \frac{\mu_X}{\mu_Y} + \frac{\mu_X}{(\mu_Y)^3} \text{Var}[Y]. \quad (A.4)$$

Using (A.3) and (A.4), one then gets

$$\text{Var} \left[ \frac{X}{Y} \right] \approx E \left[ \frac{X}{\mu_Y} - \frac{\mu_X}{(\mu_Y)^2} (Y - \mu_Y) + \frac{\mu_X}{(\mu_Y)^3} (Y - \mu_Y)^2 - \frac{1}{(\mu_Y)^2} (Y - \mu_Y) (X - \mu_X) \right]^2 - \frac{\mu_X}{\mu_Y} \text{Var}[Y]^2$$

$$= E \left[ \frac{X - \mu_X}{\mu_Y} - \frac{\mu_X}{(\mu_Y)^2} (Y - \mu_Y) + \frac{\mu_X}{(\mu_Y)^3} [(Y - \mu_Y)^2 - \text{Var}[Y]] \right] - \frac{1}{(\mu_Y)^2} (Y - \mu_Y) (X - \mu_X) \right]^2.$$

Note that in the case of $\mu_X = 0$ (which is the relevant case we consider below) we then get

$$\text{Var} \left[ \frac{X}{Y} \right] \bigg|_{\mu_X = 0} \approx \frac{1}{\mu_Y^2} \text{Var}[X] + \frac{1}{\mu_Y^4} \text{Var}[X] \text{Var}[Y], \quad (A.5)$$

where we have used the independence of $X$ and $Y$. Imposing $\mu_X = 0$ on the second expression in Appendix C in EHS yields $\text{Var} \left[ \frac{X}{Y} \right] \bigg|_{\mu_X = 0} \approx \frac{1}{\mu_Y^2} \text{Var}[X]$, which thus differs from (A.5) by the second term on the right-hand side of (A.5). This difference is the result of a mistake in EHS’ computations, but it does not affect the conclusion that the approximations may lead to the wrong conclusion about the desirability of central banker preference uncertainty.

Now consider inflation variance. By (A.1), we set $X = b\varepsilon$ (hence, $\mu_X = 0$) and $Y = a_t + b^2$ in (A.5), to obtain:

$$\text{Var} [\pi] = \frac{b^2}{(a_t + b^2)^2} \sigma^2_{\varepsilon} + \frac{b^2}{(a_t + b^2)^4} \sigma^2_{\varepsilon} \sigma^2_x. \quad (A.6)$$
Due to the abovementioned computational mistake, EHS’ equation (15) erroneously report the inflation variance when \( z = 0 \) as \( \text{Var}[\pi] = \left[ \frac{b^2}{(\pi + b^2)^2} \right] \sigma^2_\pi \), cf. the Footnote 2 in our main text. With \( z = \pi^e = 0 \), \( y = y^* + b\pi + \varepsilon \), and the output variance therefore becomes

\[
\text{Var}[y] = b^2\text{Var}[\pi] + \sigma^2_\pi + 2bE[\pi\varepsilon] \\
= b^2\text{Var}[\pi] + \sigma^2_\pi - 2b^2E\left[ \frac{1}{a_t + b^2} \varepsilon^2 \right].
\]

Apply (A.4) to the final term in this expression and use (A.6) to obtain:

\[
\text{Var}[y] = \frac{\alpha^2}{(\pi + b^2)^2}\sigma^2_\pi - 2\frac{b^2(\pi + b^2/2)}{(\pi + b^2)^4}\sigma^2_\pi\sigma^2_x. \tag{A.7}
\]

From this expression, it follows that preference uncertainty reduces \( \text{Var}[y] \), whatever the value of \( \pi \). This is wrong, as we demonstrate in Section 3.

To further illustrate the crudeness of the approximation, note that society’s expected loss can be written as:

\[
E[S] = \alpha E]\left[ \pi^2 \right] + E]\left[ (y - ky^*)^2 \right] \\
= \alpha \left\{ (\pi^e)^2 + \text{Var}[\pi] \right\} + z^2 + \text{Var}[y],
\]

which for \( z = 0 \), and thus \( \pi^e = 0 \), becomes

\[
E[S] = \alpha\text{Var}[\pi] + \text{Var}[y].
\]

Set \( \pi = \alpha \) and use (A.6) and (A.7) to find:

\[
E[S] = \frac{\alpha}{\alpha + b^2}\sigma^2_\pi - \frac{b^2}{(\alpha + b^2)^3}\sigma^2_\pi\sigma^2_x. \tag{A.8}
\]

Hence, if \( \sigma^2_x > 0 \), this expression becomes smaller than the first-best expected loss, \( \alpha\sigma^2_\pi / (\alpha + b^2) \), which can obviously not be correct.

### A.2. A more accurate approximation

One may note that if the statistical independence of \( \varepsilon \) and \( a_t \) is acknowledged before performing any approximations, then it is straightforward that one does not need to approximate \( \varepsilon \) around its mean (as EHS do), because it enters linearly in the expressions for inflation and output; cf. (A.1) and (A.2). Only the coefficients of \( \varepsilon \) need to be approximated, as these are non-linear functions of \( a_t \). Now, we derive the approximated
expressions for the inflation variance, the output variance and the expected social loss with an approximation that makes in this way use of the statistical independence of \( \varepsilon \) and \( a_t \). With this approximation one is not led to wrong conclusions.

From (A.1), it follows immediately by the independence of \( \varepsilon \) and \( a_t \) (as \( E[\varepsilon] = 0 \)) that the inflation variance is given by

\[
\text{Var} [\pi] = \mathbb{E} \left[ \frac{b}{a_t + b^2} \right]^2 \mathbb{E} [\varepsilon]^2 = \mathbb{E} \left[ \frac{1}{(a_t + b^2)^2} \right] b^2 \sigma^2 \varepsilon,
\]

which is increasing in \( \sigma^2 \varepsilon \) because the term in the square brackets is strictly convex in \( a_t \) (cf. Footnote 2 in the main text). We then perform a second-order Taylor approximation on the term in the square brackets around \( a_t = \bar{a} \):

\[
\frac{1}{(a_t + b^2)^2} \simeq \frac{1}{(\bar{a} + b^2)^2} - \frac{2}{(\bar{a} + b^2)^3} (a_t - \bar{a}) + \frac{3}{(\bar{a} + b^2)^4} (a_t - \bar{a})^2.
\]

Taking expectations over this, and inserting the result back into the above expression, we get the following approximation for inflation variance:

\[
\text{Var} [\pi] \simeq \frac{b^2}{(\bar{a} + b^2)^2} \sigma^2 \varepsilon + \frac{3b^2}{(\bar{a} + b^2)^4} \sigma^2 \varepsilon^2.
\]

Note that with this approximation, the coefficient of \( \sigma^2 \varepsilon^2 \) in the expression for \( \text{Var}[\pi] \) is three times larger than the corresponding coefficient in EHS’ approximation — compare (A.11) with (A.6). This confirms what is stated in Section 3 of the main text. From (A.2), it follows that the output variance is given by

\[
\text{Var} [y] = \mathbb{E} \left[ \frac{a_t}{a_t + b^2} \right]^2 \mathbb{E} [\varepsilon]^2 = \mathbb{E} \left[ \frac{a_t^2}{(a_t + b^2)^2} \right] \sigma^2 \varepsilon.
\]

We then perform a second-order Taylor approximation on the term in the square brackets around \( a_t = \bar{a} \):

\[
\frac{a_t^2}{(a_t + b^2)^2} \simeq \frac{\bar{a}^2}{(\bar{a} + b^2)^2} + \frac{2\bar{a}b^2}{(\bar{a} + b^2)^3} (a_t - \bar{a}) + \frac{b^2 (b^2 - 2\bar{a})}{(\bar{a} + b^2)^4} (a_t - \bar{a})^2.
\]

Taking expectations over this, and inserting the result back into the above expression, we
get the following approximation for output variance:

\[
\text{Var} [y] \simeq \frac{\bar{\alpha}^2}{(\bar{\alpha} + b^2)^2} \sigma_e^2 + \frac{b^2 (b^2 - 2 \bar{\alpha})}{(\bar{\alpha} + b^2)^4} \sigma_e^2 \sigma_x^2.
\]  

(A.13)

In contrast to (A.7), we see that preference uncertainty does not unambiguously reduce output variance. Even if it does reduce output variance, the reduction is smaller (see the final paragraph of Section 3 in the main text) than the reduction obtained under EHS’ approximation when the latter is correctly computed — compare (A.13) with (A.7). The output variance reduction under the correctly computed EHS approximation, in turn, is smaller than the reduction in output variance reported by EHS, which is based on the abovementioned incorrectly computed inflation variance — see EHS’ equation (16) for \( z = 0 \). Equation (A.13) is in conformity with what we established in the main text by use of the exact expression for the output variance (and note that, as we explained in Footnote 1 in the main text, output variance is likely to increase with preference uncertainty if \( \bar{\alpha} \) is relative low, and \( b \) is relatively high). Indeed, a sufficient condition for preference uncertainty to raise output variance is shown in Appendix B to be \( 2 (\bar{\alpha} - x_t) < b^2 \) for all possible realizations of \( x_t \) [see equation (B.2)]. This is exactly the condition one derives from (A.13) when acknowledging that the approximation is performed around \( a_t = \bar{\alpha} \), i.e., \( x_t = 0 \).

Finally, consider the expected social loss, where, as above, we examine the case of \( \alpha = \bar{\alpha} \). Since \( E[S] = \alpha \text{Var}[\pi] + \text{Var}[y] \), cf. above, we immediately get from (A.11) and (A.13) that

\[
E[S] \simeq \frac{\alpha}{\alpha + b^2} \sigma_e^2 + \frac{b^2}{(\alpha + b^2)^3} \sigma_e^2 \sigma_x^2.
\]  

(A.14)

In contrast to (A.8), it here follows that preference uncertainty is harmful, as it unambiguously increases the expected social loss in (A.14) [and correctly implies that the expected social loss exceeds the first-best expected loss — the first term on the right-hand side of (A.14)].

A.3. Demonstration that the last approximation is more accurate

We now demonstrate why the last approximation is more accurate. Consider the function \( G (X, Y) = f (X) g (Y) \). By definition, the best second-order approximation of \( G \) around
$X = \mu_X$ and $Y = \mu_Y$ is the standard Taylor expression:

$$G(X,Y) \approx f(\mu_X) g(\mu_Y) + f'(\mu_X) g(\mu_Y)(X - \mu_X) + f(\mu_X) g'(\mu_Y)(Y - \mu_Y)$$

$$+ \frac{1}{2} f''(\mu_X) g(\mu_Y)(X - \mu_X)^2 + \frac{1}{2} f(\mu_X) g''(\mu_Y) (Y - \mu_Y)^2$$

$$+ f'(\mu_X) g'(\mu_Y)(X - \mu_X)(Y - \mu_Y).$$  \hspace{1cm} (A.15)

In the case of EHS, $f(X) = X$ and $g(Y) = 1/Y$. Substituting these expressions into (A.15) gives (A.3).

The approximations for $\pi$ and $y$ in the previous subsection, on the other hand, are based on the following approximation of $G$:

$$G(X,Y) \approx f(X) \left[ g(\mu_Y) + g'(\mu_Y)(Y - \mu_Y) + \frac{1}{2} g''(\mu_Y)(Y - \mu_Y)^2 \right].$$  \hspace{1cm} (A.16)

In our case, this is a more accurate approximation than (A.15) as we now show. Start by making a third-order Taylor approximation of $G$ around $X = \mu_X$ and $Y = \mu_Y$:

$$G(X,Y) \approx f(\mu_X) g(\mu_Y) + f'(\mu_X) g(\mu_Y)(X - \mu_X) + f(\mu_X) g'(\mu_Y)(Y - \mu_Y)$$

$$+ \frac{1}{2} f''(\mu_X) g(\mu_Y)(X - \mu_X)^2 + \frac{1}{2} f(\mu_X) g''(\mu_Y) (Y - \mu_Y)^2$$

$$+ f'(\mu_X) g'(\mu_Y)(X - \mu_X)(Y - \mu_Y)$$

$$+ \frac{1}{6} f'''(\mu_X) g(\mu_Y)(X - \mu_X)^3 + \frac{1}{6} f(\mu_X) g'''(\mu_Y) (Y - \mu_Y)^3$$

$$+ \frac{3}{6} f''(\mu_X) g'(\mu_Y)(X - \mu_X)^2(Y - \mu_Y)$$

$$+ \frac{3}{6} f(\mu_X) g''(\mu_Y)(X - \mu_X)(Y - \mu_Y)^2$$  \hspace{1cm} (A.17)

Now use that for both $\pi$ and $y$ we have $f(X) = X$ and $\mu_X = 0$ (as $X$ equals the supply shock $\varepsilon$). Hence, (A.16) becomes

$$G(X,Y) \approx X \left[ g(\mu_Y) + g'(\mu_Y)(Y - \mu_Y) + \frac{1}{2} g''(\mu_Y)(Y - \mu_Y)^2 \right],$$  \hspace{1cm} (A.18)
and (A.17) becomes:

\[
G(X, Y) \simeq 0 + g(\mu_Y)X + 0 \\
+ \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 0 \\
+ g'(\mu_Y)X(Y - \mu_Y) \\
+ \frac{1}{6} \cdot 0 + \frac{1}{6} \cdot 0 \\
+ \frac{3}{6} \cdot 0 \\
+ \frac{3}{6} g''(\mu_Y)X(Y - \mu_Y)^2 \\
= X \left[ g(\mu_Y) + g'(\mu_Y)(Y - \mu_Y) + \frac{1}{2} g''(\mu_Y)(Y - \mu_Y)^2 \right].
\]

which is the same as (A.18). Hence, the approximation used in the previous subsection is a third-order Taylor approximation of \( G \) around \((0, \mu_Y)\). While a third-order approximation yields qualitatively correct conclusions about the effects of preference uncertainty, EHS’ second-order Taylor approximations are obviously too crude as they lead to qualitatively incorrect conclusions concerning the effects of preference uncertainty.

### B. Condition for convexity

The solution for output is \( y = y^* + (a_t/[a_t + b^2]) \varepsilon \). Hence, using the independence of \( x_t \) and \( \varepsilon \), we have:

\[
\text{Var}[y] = \mathbb{E} \left( \frac{(\alpha - x_t)^2}{(\alpha - x_t + b^2)^2} \varepsilon^2 \right) \\
= \mathbb{E} \left[ \frac{(\alpha - x_t)^2}{(\alpha - x_t + b^2)^2} \right] \sigma^2_{\varepsilon}. \tag{B.1}
\]

We want to see under what circumstances the term in square brackets is strictly convex in \( x_t \). The first-order derivative of this term with respect to \( x_t \) is:

\[
\frac{-2(\alpha - x_t)(\alpha - x_t + b^2)^2 + 2(\alpha - x_t + b^2)(\alpha - x_t)^2}{(\alpha - x_t + b^2)^4} \\
= \frac{-2(\alpha - x_t)(\alpha - x_t + b^2) + 2(\alpha - x_t)^2}{(\alpha - x_t + b^2)^3} \\
= \frac{-2(\alpha - x_t) b^2}{(\alpha - x_t + b^2)^3}.
\]
The second-order derivative is therefore
\[
\frac{2b^2 (\bar{a} - x_t + b^2)^3 - 6 (\bar{a} - x_t + b^2) (\bar{a} - x_t) b^2}{(\bar{a} - x_t + b^2)^6} = \frac{2b^2 (\bar{a} - x_t + b^2) - 6 (\bar{a} - x_t) b^2}{(\bar{a} - x_t + b^2)^4} = \frac{2b^2 [b^2 - 2 (\bar{a} - x_t)]}{(\bar{a} - x_t + b^2)^4}.
\]

Hence, the term in square brackets in (B.1) is strictly convex in \(x_t\) if
\[
2 (\bar{a} - x_t) < b^2. \tag{B.2}
\]

A sufficient (though not necessary) condition for preference uncertainty to raise \(\text{Var}[y]\) is that the term in square brackets in (B.1) is strictly convex in \(x_t\) for all possible realizations of \(x_t\). As (B.2) shows, if \(x_t > -\bar{a}\) with probability 1, then this is the case when \(4\bar{a} < b^2\).

C. Derivation of expected social loss with a two-point preference uncertainty distribution

First note that society’s expected loss can be written as
\[
E[S] = \alpha \{ (\pi^e)^2 + \text{Var} [\pi] \} + z^2 + \text{Var} [y]. \tag{C.1}
\]

The distribution of shocks to the central banker’s preferences is given by:
\[
a_t = \begin{cases} 
\bar{a} + \Delta, & \text{with probability } \frac{1}{2}, \\
\bar{a} - \Delta, & \text{with probability } \frac{1}{2}.
\end{cases} \tag{C.2}
\]

Using (5), the rationality of expectations and the independence of \(a_t\) and \(\varepsilon\), we find that
\[
\pi^e = \left[ \frac{E \left( \frac{b}{a_t + b^2} \right)}{1 - E \left( \frac{b^2}{a_t + b^2} \right)} \right] z.
\]

Further, using (C.2) we find that
\[
E \left( \frac{b}{a_t + b^2} \right) = \left[ \frac{\bar{a} + b^2}{(\bar{a} + b^2)^2 - \Delta^2} \right] b.
\]
\[ \pi^e = \left[ \frac{\pi + b^2}{(\pi + b^2)\pi - \Delta^2} \right] b^2. \]  

(C.3)

Insert (5) for \( \pi \) in (1), which gives:

\[ y = y^* + \left( -\frac{a_t}{a_t + b^2} \pi^e + \frac{b}{a_t + b^2} z \right) + \left( \frac{a_t}{a_t + b^2} \varepsilon. \right) \]

Of course, the expectation of the term in square brackets is zero. One has therefore that output variance is given by:

\[ \text{Var} \left[ y \right] = \varepsilon \left( -\left( \frac{a_t}{a_t + b^2} \right) \pi^e + \left( \frac{b}{a_t + b^2} \right) z \right)^2 + \varepsilon \left( \frac{a_t}{a_t + b^2} \right)^2. \]  

(C.4)

Further, one has that inflation variance is given by

\[ \text{Var} \left[ \pi \right] = \varepsilon \left[ \left( \frac{b}{a_t + b^2} \right) \left( b\pi^e + z - \varepsilon - \pi^e \right) \right]^2 \]

\[ = \varepsilon \left[ \left( \frac{b}{a_t + b^2} \right) \pi^e + \left( \frac{b}{a_t + b^2} \right) z \right]^2 + \varepsilon \left[ \left( \frac{b}{a_t + b^2} \right)^2 \varepsilon^2 \right]. \]  

(C.5)

By (C.2), it follows that

\[ \varepsilon \left[ \left( \frac{a_t}{a_t + b^2} \right) \pi^e + \left( \frac{b}{a_t + b^2} \right) z \right]^2 \]

\[ = \frac{1}{2} \left[ \left( \frac{a_t - \Delta}{a_t - \Delta + b^2} \right) \pi^e + \left( \frac{b}{a_t + b^2} \right) z \right]^2 + \frac{1}{2} \left[ \left( \frac{\pi + \Delta}{a_t - \Delta + b^2} \right) \pi^e + \left( \frac{b}{a_t + b^2} \right) z \right]^2. \]  

(C.6)

We can now write down the expected social loss, by substituting the right-hand sides of (C.4) and (C.5) into (C.1):

\[ \text{E}[S] = (\alpha + b^2) \varepsilon \left[ -\left( \frac{a_t}{a_t + b^2} \pi^e + \left( \frac{b}{a_t + b^2} \right) z \right)^2 + z^2 + \alpha (\pi^e)^2 \right] \]

\[ + \frac{\alpha b^2 \left[ (\pi + b^2)^2 + \Delta^2 \right] + \pi^2 (\pi + b^2)^2 + \Delta^2 \left[ \Delta^2 + b^4 - 2\pi (\pi + b^2) \right]}{[\pi + b^2]^2 - \Delta^2}] \varepsilon^2, \]  

(C.7)

where the final term is obtained by applying (C.2) to

\[ \alpha \varepsilon \left[ \left( \frac{b}{a_t + b^2} \right)^2 \varepsilon^2 \right] + \varepsilon \left[ \left( \frac{a_t}{a_t + b^2} \right)^2 \varepsilon^2 \right]. \]
Numerical expressions for $E[S]$ are thus obtained by substituting (C.3) into (C.6) and substituting the result and, again, (C.3) into (C.7).

D. Proof that, for given $\pi$, preference uncertainty may reduce the expected welfare loss

We provide an example of a case in which some preference uncertainty reduces the expected welfare loss. Take the case of the two-point distribution in Section 4. Now set $z = 0$. Hence,

$$E[S] = \frac{\alpha b^2 \left[ (\bar{\pi} + b^2)^2 + \Delta^2 \right] + \bar{\pi}^2 (\bar{\pi} + b^2)^2 + \Delta^2 \left[ \Delta^2 + b^4 - 2\bar{\pi} (\bar{\pi} + b^2) \right]}{\left[ (\bar{\pi} + b^2)^2 - \Delta^2 \right]^2} \sigma^2_\varepsilon.$$

Note that if $\bar{\pi} = \alpha$ and $\Delta = 0$, then this expression reduces to the first-best expected loss. Now, take $\alpha = b = \sigma^2_\varepsilon = 1$ and $\bar{\pi} = 5$. If $\Delta = 0$ (no preference uncertainty), then $E[S] = (36 + 25 \times 36)/36^2 = 0.722$. If $\Delta = 1$, then $E[S] = (37 + 25 \times 36 - 58)/35^2 = 0.718$, which confirms the possibility that preference uncertainty can reduce the expected welfare loss, for given $\bar{\pi}$.

E. Proof that the inflation and output variances increase for any $b > 0$ when the loss function of Footnote 7 applies

In the case of an arbitrary $b > 0$, the loss function is, according to Footnote 7:

$$L = a_t \pi^2 + (1 + \bar{\pi} - a_t) (y - ky^*)^2 / b^2.$$

It is straightforward to find the reaction function as

$$\pi = \left[ \frac{1 + \bar{\pi} - a_t}{b (1 + \bar{\pi})} \right] (b\pi e + z - \varepsilon).$$
Inflation expectations then follow as $E[\pi] = \pi_e = z / (b\pi)$.\(^2\) To find the solutions for inflation and output, substitute this expression for $\pi_e$ into the central bank reaction function to obtain:

$$\pi = \frac{1 + a - a_t}{b\alpha} z - \frac{1 + a - a_t}{b(1 + \alpha)} \varepsilon$$

$$= \frac{1 + x_t}{b\alpha} z - \frac{1 + x_t}{b(1 + \alpha)} \varepsilon,$$

from which output follows as

$$y = y^* + \frac{\alpha - a_t}{\alpha} z + \frac{a_t}{1 + \alpha} \varepsilon$$

$$= y^* + \frac{x_t}{\alpha} z + \frac{\alpha - x_t}{1 + \alpha} \varepsilon.$$

Using the independence of $x_t$ and $\varepsilon$, we have that:

$$\text{Var}[\pi] = \left(\frac{z}{b\alpha}\right)^2 \sigma_x^2 + \frac{1}{b^2 (1 + \alpha)^2} (1 + \sigma_x^2) \sigma_\varepsilon^2,$$

$$\text{Var}[y] = \left(\frac{z}{\alpha}\right)^2 \sigma_x^2 + \frac{1}{(1 + \alpha)^2} (\alpha^2 + \sigma_x^2) \sigma_\varepsilon^2,$$

which are both increasing in $\sigma_x^2$. Hence, given that the expected values of $\pi$ and $y$ do not depend on $x_t$, the expected social loss is increasing in $\sigma_x^2$.

---

\(^2\)This suggests that the inflation bias is decreasing with $b$, i.e., the output gain from surprise inflation, which is in contrast with the usual intuition. The result arises because a higher $b$ also reduces the weight on the output target in the loss function. Hence, when one examines the equilibrium implications of raising $b$, one must keep the weight on output relative to inflation constant. Formally, under certainty (i.e., $a_t = \pi$), an increase in $b$ reduces the weight on output stabilization by $2(\db/b)$ percent. Hence, to keep the relative weight constant, $\pi$ must fall by the same proportion. That is, $\text{d}\pi/\pi = -2\db/b$ must hold. To perform a proper analysis of how more output gains of inflation surprises affect the inflation bias, use the expression for inflation expectations to find:

$$d\pi_e = -\frac{z}{b\alpha} \frac{db}{b} - \frac{z}{b\alpha} \frac{d\pi}{\alpha}.$$

Then use the above requirement on $d\pi/\pi$ to find that

$$d\pi_e = \frac{z}{b\alpha} \frac{db}{b} > 0.$$

That is, for an unchanged relative weight of the central bank’s objectives, the model confirms the familiar result that a higher gain from surprise inflation (i.e., a higher $b$) leads to a higher inflation bias.