Online Appendix: Forward Guidance without Common Knowledge

George-Marios Angeletos and Chen Lian*

July 22, 2018

Appendix A: Proofs

In this Appendix, we prove the results stated in the main text. For all the proofs that regard the New-Keynesian model (as opposed to the abstract analysis in Section IV), we use a tilde over a variable to denote the log-deviation of this variable from its steady-state counterpart, and reserve the non-tilde notation for the original variables. The only exception to this rule is that we let $\tilde{a}_{i,t} \equiv \frac{a_{i,t}}{c^*}$, where $a_{i,t}$ is consumer i's initial asset position at period-t and c^* is steady-state spending. This takes care of the issue that the log-deviation of the asset position is not well defined because the steady-state value is $a^* = 0$ and is standard in the literature (e.g., Woodford, 2011).

Proof of Proposition 1. We proceed in four steps, starting with the behavior of the consumers, proceeding with the behavior of the firms, and concluding with market clearing and with the derivation of the two beauty contests shown in the main text.

Step 1: Consumers. Consider an arbitrary consumer $i \in \mathcal{I}_c$. Let $a_{i,t} = R_{t-1}s_{i,t-1}/\pi_t$ denote consumer i's initial asset position at period-t. By condition (2), the following intertemporal budget constraint holds in all periods and all states of Nature:[1]

$$\sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\tfrac{R_{t+j-1}}{\pi_{t+j}}\right)^{-1}\right]c_{i,t+k}\right\} = a_{i,t}+\sum_{k=0}^{+\infty}\left\{\left[\prod_{j=1}^{k}\left(\tfrac{R_{t+j-1}}{\pi_{t+j}}\right)^{-1}\right](w_{i,t+k}n_{i,t+k}+e_{i,t+k})\right\} \quad . \tag{26}$$

*Angeletos: Department of Economics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 (email:angelet@mit.edu); Lian: Department of Economics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 (email:lianchen@mit.edu).

[1] One should think of the state of Nature as a realization of the exogenous payoff relevant shocks along with the cross-sectional distribution of the exogenous signals (information) received by the agents.

Taking the log-linear approximation of the above around the steady state, we get the following:

$$\sum_{k=0}^{+\infty} \beta^k \tilde{c}_{i,t+k} = \tilde{a}_{i,t} + \sum_{k=0}^{+\infty} \beta^k \left\{ \Omega \left(\tilde{w}_{i,t+k} + \tilde{n}_{i,t+k} \right) + (1-\Omega)\, \tilde{e}_{i,t+k} \right\}, \tag{27}$$

where Ω is the ratio of labor income to total income in steady state. The consumer's optimality conditions, on the other hand, can be expressed as follows:

$$\tilde{n}_{i,t} = \frac{1}{\epsilon} \left(\tilde{w}_{i,t} - \frac{1}{\sigma} \tilde{c}_{i,t} \right), \tag{28}$$

$$\tilde{c}_{i,t} = E_{i,t} \left[\tilde{c}_{i,t+1} - \sigma \left(\tilde{R}_t - \tilde{\pi}_{t+1} \right) \right] = E_{i,t} \left[\tilde{c}_{i,t+1} - \sigma \tilde{r}_{t+1} \right], \tag{29}$$

where $E_{i,t}[\cdot]$ is the expectation of consumer i in period t. The first condition describes optimal labor supply; the second is the *individual-level* Euler condition, which describes optimal consumption and saving.

At this point, it is worth emphasizing that our analysis preserves the *standard* Euler condition at the individual level. This contrasts with McKay, Nakamura and Steinsson (2016, 2017) and Werning (2015), where liquidity constraints cause this condition to be violated for *some* agents, as well as with Gabaix (2016), where a cognitive friction causes this condition to be violated for *every* agent. We revisit this point in Appendix C, when we show that our analysis rationalizes a *discounted* Euler condition at the aggregate level, *in spite of* the preservation of the standard condition at the individual level.

Combining conditions (27), (28) and (29), we obtain the optimal expenditure of consumer i in period t as a function of the current and the expected future values of wages, dividends, and real interest rates:

$$\begin{aligned} \tilde{c}_{i,t} = {} & \tfrac{(1-\beta)\epsilon\sigma}{\epsilon\sigma+\Omega} \tilde{a}_{i,t} - \sigma \sum_{k=1}^{+\infty} \beta^k E_{i,t} \left[\tilde{r}_{t+k} \right] \\ & + (1-\beta) \left[\tfrac{(\epsilon+1)\sigma\Omega}{\epsilon\sigma+\Omega} \tilde{w}_{i,t} + \tfrac{\epsilon\sigma(1-\Omega)}{\epsilon\sigma+\Omega} \tilde{e}_{i,t} \right] + (1-\beta) \sum_{k=1}^{+\infty} \beta^k E_{i,t} \left[\tfrac{(\epsilon+1)\sigma\Omega}{\epsilon\sigma+\Omega} \tilde{w}_{i,t+k} + \tfrac{\epsilon\sigma(1-\Omega)}{\epsilon\sigma+\Omega} \tilde{e}_{i,t+k} \right]. \end{aligned} \tag{30}$$

This condition, which is a variant of the consumption function seen in textbook treatments of the Permanent Income Hypothesis,[2] contains two elementary insights. First, all future variables—wages, dividends, and real interest rates—are discounted. Second, the *current* spending of a consumer depends on the present value of her income, which in turn depends, in equilibrium, on the *future* spending of other consumers.

The first property guarantees that the decision-theoretic, or partial-equilibrium, effect of forward guidance diminishes with the horizon at which interest rates are changed; the second represents a dynamic strategic complementarity, which is the modern reincarnation what was known as the "income multiplier"

[2] To see this more clearly, suppose that initial assets are zero, that the real interest rate is expected to equal the discount rate at all periods, and that labor supply is fixed ($\epsilon \to \infty$). Condition (30) then reduces to $\tilde{c}_{i,t} = (1-\beta)\left[\Omega \tilde{w}_{i,t} + (1-\Omega)\, \tilde{e}_{i,t}\right] + (1-\beta) \sum_{k=1}^{+\infty} \beta^k E_{i,t} \left[\Omega \tilde{w}_{i,t+k} + (1-\Omega)\, \tilde{e}_{i,t+k}\right]$, which means that optimal consumption equals "permanent income" (the annuity value of current and future income). Relative to this benchmark, condition (30) adjusts for three factors: for the endogeneity of labor supply, which explains the different weights on wages and dividends; for initial assets, which explains the first term in condition (30); and for the potential gap between the real interest rate and the subjective discount rate, which explains the second term.

in the IS-LM framework. We elaborate on these two points more in the main text. For the time being, we aggregate condition (30), and use the facts that assets average to zero and that future idiosyncratic shocks are unpredictable, to obtain the following condition for aggregate spending:

$$\tilde{c}_t = -\sigma \sum_{k=1}^{+\infty} \beta^k \bar{E}_t^c \left[\tilde{r}_{t+k}\right] + (1-\beta)\left[\tfrac{(\epsilon+1)\sigma\Omega}{\epsilon\sigma+\Omega}\tilde{w}_t + \tfrac{\epsilon\sigma(1-\Omega)}{\epsilon\sigma+\Omega}\tilde{e}_t\right] \tag{31}$$
$$+ (1-\beta)\sum_{k=1}^{+\infty} \beta^k \bar{E}_t^c \left[\tfrac{(\epsilon+1)\sigma\Omega}{\epsilon\sigma+\Omega}\tilde{w}_{t+k} + \tfrac{\epsilon\sigma(1-\Omega)}{\epsilon\sigma+\Omega}\tilde{e}_{t+k}\right],$$

where $\bar{E}_t^c\left[\cdot\right]$ henceforth denotes the *average* expectation of the consumers in period t.

Step 2: Firms. Consider a firm $j \in \mathcal{I}_f$ that gets the chance to reset its price during period t. The optimal reset price, denoted by p_t^{j*}, is given by the following:

$$\tilde{p}_t^{j*} = (1-\beta\theta)\left\{(\tilde{mc}_t^j + \tilde{p}_t) + \sum_{k=1}^{+\infty} (\beta\theta)^k E_{j,t}\left[\tilde{mc}_{t+k}^j + \tilde{p}_{t+k}\right]\right\} + (1-\beta\theta)\,\tilde{\mu}_t^j, \tag{32}$$

where $E_{j,t}^f\left[\cdot\right]$ denotes the firm's expectations in period t, $\tilde{mc}_t^j = \tilde{w}_t^j$ is its *real* marginal cost in period t, and $\tilde{\mu}_t^j$ is the corresponding markup shock. The interpretation of this condition is familiar: the optimal "reset" price is given by the expected nominal marginal cost over the expected lifespan of the new price, plus the markup.[3] Aggregating the above condition, using the fact that the past price level is known and that inflation is given by $\tilde{\pi}_t = (1-\theta)\left(\tilde{p}_t^* - \tilde{p}_{t-1}\right)$, where $\tilde{p}_t^* \equiv \int_{\mathcal{I}_f} \tilde{p}_t^{j*} dj$, we obtain the following condition for the level of inflation in period t:

$$\tilde{\pi}_t = \varkappa \tilde{mc}_t + \varkappa \sum_{k=1}^{+\infty} (\beta\theta)^k \bar{E}_t^f[\tilde{mc}_{t+k}] + \tfrac{1-\theta}{\theta}\sum_{k=1}^{+\infty} (\beta\theta)^k \bar{E}_t^f[\tilde{\pi}_{t+k}] + \varkappa\tilde{\mu}_t, \tag{33}$$

where $\varkappa \equiv \frac{(1-\theta)(1-\beta\theta)}{\theta}$ and $\bar{E}_t^f\left[\cdot\right]$ henceforth denotes the average expectation of the firms. The latter may or may not be the same as the average expectation of the consumers.

Step 3: Market Clearing, Wages, and Profits. Because the final-good sector is competitive and observes all the relevant prices,[4] and because the technology satisfies (3) and (4), we have that $\tilde{p}_t = \int_{\mathcal{I}_f} \tilde{p}_t^j dj$ and $\tilde{y}_t = \int_{\mathcal{I}_f} \tilde{y}_t^j dj = \int_{\mathcal{I}_f} \tilde{l}_t^j dj$. The latter, together with market clearing in the labor market, gives $\tilde{y}_t = \tilde{n}_t \equiv \int_{\mathcal{I}_c} \tilde{n}_{i,t} di$. Market clearing in the market for the final good, on the other hand, gives

$$\tilde{y}_t = \tilde{c}_t \equiv \int_{\mathcal{I}_c} \tilde{c}_{i,t} di.$$

Finally, note that the real profit of monopolist j at period t is given by $e_t^j = \left(\frac{p_t^j}{p_t} - w_t^j\right) y_t^j$. Log-linearizing and aggregating it gives $\tilde{e}_t = -\frac{\Omega}{1-\Omega}\tilde{w}_t + \tilde{y}_t$. Combining all these facts with (28), the optimality condition for

[3] Note that future markups are unpredictable.

[4] Recall that the we have allowed the entire price vector, $(p_t^j)_{j\in[0,1]}$, to be common knowledge at period t.

labor supply, we arrive at the following characterization of the aggregate wages and the profits:

$$\tilde{w}_t = \tilde{mc}_t = \left(\epsilon + \frac{1}{\sigma}\right)\tilde{y}_t, \qquad \tilde{e}_t = \left[1 - \frac{\Omega\left(\epsilon + \frac{1}{\sigma}\right)}{1-\Omega}\right]\tilde{y}_t, \quad \text{and} \quad \frac{(\epsilon+1)\sigma\Omega}{\epsilon\sigma+\Omega}\tilde{w}_t + \frac{\epsilon\sigma(1-\Omega)}{\epsilon\sigma+\Omega}\tilde{e}_t = \tilde{y}_t. \tag{34}$$

Step 4: Beauty Contests. Condition (31), which follows merely from consumer optimality, pins down aggregate spending as a function of the average beliefs of wages, profits, interest rates, and inflation. As we impose REE, a consumer can infer that (34) holds, aggregate spending can then be expressed as a function of the consumers' average beliefs of interest rates, of inflation, and of aggregate spending itself. This is condition (5), the consumption beauty contest. Similarly, combining (33) and (34), we can express aggregate inflation as a function of the firms' average beliefs of aggregate spending and of inflation itself. This is condition (6), the inflation beauty contest.

Proof of Proposition 2. Because Θ_t is zero for all $t > T$, a_t is also zero for all $t > T$.[5] Using this fact along with the fact that Θ_t is zero also for $t < T$, and iterating on condition (12), we can obtain a_t for all $t < T$ as a linear function of the average first- and higher-order beliefs about Θ_T; see, e.g., Lemma 2 below for an explicit characterization in the case without learning. When information is complete, all agents share the same first-order beliefs about Θ_T with probability one, and this fact is itself common knowledge. It follows that higher-order beliefs collapse to first-order beliefs and, therefore, a_t becomes a linear function of $E_t[\Theta_T]$, the commonly shared expectation of Θ_T. Now take any $t < \tau \leq T$ and any pair of agents i, j. Complete information guarantees that $E_{i,t}[E_\tau[\Theta_T]] = E_t[\Theta_T] = E_{j,t}[E_\tau[\Theta_T]]$ with probability one. And since we already argued that, in equilibrium, a_τ is a known linear function of $E_\tau[\Theta_T]$, it is also the case $E_{i,t}[a_\tau] = E_{j,t}[a_\tau]$. That is, complete information (in the sense of Definition 1) rules out imperfect consensus (in the sense of Definition 2).

Proof of Lemma 1. Lemma 1 directly follows from the argument in main text.

Proof of Lemma 2. We prove the following stronger result: there exists positively-valued coefficients $\{\chi_{h,k}\}_{k\geq 1, 1\leq h\leq k}$, such that, for any $t \leq T-1$,

$$a_t = \sum_{h=1}^{T-t}\left\{\chi_{h,T-t}\bar{E}_t^h\left[\Theta_T\right]\right\}, \tag{35}$$

where each $\chi_{h,k}$ is a function of (α, γ, h, k) and $\bar{E}_t^h[\cdot]$ is defined recursively by $\bar{E}_t^1[\cdot] = \bar{E}_t[\cdot]$ and $\bar{E}_t^h[\cdot] = \bar{E}_t\left[\bar{E}_t^{h-1}[\cdot]\right]$ for every $h \geq 2$. We now prove this claim by induction. First, consider $t = T-1$. From

[5] As mentioned in main text, we assume $\lim_{k\to\infty}\gamma^k E_{i,t}\left[a_{t+k}\right] = 0$ and rule out "extrinsic bubbles."

$a_T = \Theta_T$[6] and condition (16), we have $a_{T-1} = (\gamma + \alpha)\, \bar{E}_{T-1}[\Theta_T]$. It follows that condition (35) holds for

$$\chi_{1,1} = \gamma + \alpha.$$

Now, pick an arbitrary $t \leq T-2$, assume that condition (35) holds for all $\tau \in \{t+1, ..., T-1\}$, and let us prove that it also holds for t. From condition (16), we have

$$\begin{aligned} a_t &= \gamma^{T-t-1}(\gamma+\alpha)\, \bar{E}_t[\Theta_T] + \alpha \sum_{k=1}^{T-t-1} \gamma^{k-1} \bar{E}_t \left[\sum_{h=1}^{T-t-k} \left\{ \chi_{h,T-t-k} \bar{E}_{t+k}^{h}[\Theta_T] \right\} \right] \\ &= \gamma^{T-t-1}(\gamma+\alpha)\, \bar{E}_t[\Theta_T] + \sum_{h=1}^{T-t-1} \sum_{k=1}^{T-t-h} \left(\alpha \gamma^{k-1} \chi_{h,T-t-k} \right) \bar{E}_t^{h+1}[\Theta_T], \end{aligned} \tag{36}$$

where the second line uses Assumption 1 (no learning). As a result, condition (35) holds for

$$\chi_{1,T-t} = \gamma^{T-t-1}(\gamma+\alpha) \quad \text{and} \quad \chi_{h+1,T-t} = \sum_{k=1}^{T-t-h} \alpha \gamma^{k-1} \chi_{h,T-t-k} \;\; h \in \{1, \cdots T-t-1\}. \tag{37}$$

This finishes the proof.

Proof of Theorem 1. This theorem builds on Proposition 3 and Theorem 2, which are proved in the sequel. We invite the reader to read first the proofs of these two results. Here, we prove Theorem 1 taking for granted these results.

Part (i) follows directly from projecting a_0 on $\bar{E}_0[\Theta_T]$ and letting ϕ_T be the coefficient of this projection and ϵ the residual.

To prove part (ii), note that from Lemma 2, we have $\phi_T = \sum_{h=1}^{T} \chi_{h,T} \beta_h$, which is condition (20) in the main text. Together with the expression of ϕ_T^*, condition (18), and the fact that $\beta_h < 1$ for all $h \geq 2$ (from Proposition 3), we have $\phi_T / \phi_T^* < 1$ for all $T \geq 2$.

To prove part (iii), from condition (20), we have $\phi_T/\phi_T^* = \left[\sum_{h=1}^{T-1} s_{h,T}(\beta_h - \beta_{h+1}) + s_{T,T}\beta_T \right] / s_{T,T}$ and $\phi_{T+1}/\phi_{T+1}^* = \left[\sum_{h=1}^{T-1} s_{h,T+1}(\beta_h - \beta_{h+1}) + s_{T,T+1}(\beta_T - \beta_{T+1}) + s_{T+1,T+1}\beta_{T+1} \right] / s_{T+1,T+1}$. From Proposition 3 we know $\beta_h > \beta_{h+1}$ for all h. Together with Theorem 2, we have, for all $T \geq 1$,

$$\begin{aligned} \phi_{T+1}/\phi_{T+1}^* &< \left[\sum_{h=1}^{T-1} s_{h,T+1}(\beta_h - \beta_{h+1}) + s_{T+1,T+1}(\beta_T - \beta_{T+1}) + s_{T+1,T+1}\beta_{T+1} \right] / s_{T+1,T+1} \\ &= \left[\sum_{h=1}^{T-1} s_{h,T+1}(\beta_h - \beta_{h+1}) + s_{T+1,T+1}\beta_T \right] / s_{T+1,T+1} \\ &\leq \left[\sum_{h=1}^{T-1} s_{h,T}(\beta_h - \beta_{h+1}) + s_{T,T}\beta_T \right] / s_{T,T} = \phi_T/\phi_T^*. \end{aligned}$$

[6] As mentioned in main text, we assumelim$_{k\to\infty} \gamma^k E_{i,t}[a_{t+k}] = 0$ and rule out "extrinsic bubbles." Together with the fact Θ_t is zero for all $t > T$, a_t is also zero for all $t > T$. As a result, $a_T = \Theta_T$ from condition (12).

Proof of Proposition 3. Note that every agent i's signal at period 0 is drawn i.i.d. from the aggregate state of the Nature (for simplicity, we call this property "symmetry" in the rest of this proof), we have, for all $h \geq 2$,

$$\begin{aligned} Cov\left(\bar{E}_0^h[\Theta_T], \bar{E}_0^1[\Theta_T]\right) &= Cov\left(E_{i,0}\left[\bar{E}_0^{h-1}[\Theta_T]\right], \bar{E}_0^1[\Theta_T]\right) = Cov\left(E_{i,0}\left[\bar{E}_0^{h-1}[\Theta_T]\right], E_{i,0}\left[\bar{E}_0^1[\Theta_T]\right]\right), \\ &= Cov\left(\bar{E}_0^{h-1}[\Theta_T], E_{i,0}\left[\bar{E}_0^1[\Theta_T]\right]\right) = Cov\left(\bar{E}_0^{h-1}[\Theta_T], \bar{E}_0^2[\Theta_T]\right), \end{aligned}$$

where the second and the third equality come from the law of iterated expectations. By the same argument, we have, for all $h \geq 2$ and $j \in \{1, 2, \cdots h-1\}$,

$$Cov\left(\bar{E}_0^h[\Theta_T], \bar{E}_0^1[\Theta_T]\right) = Cov\left(\bar{E}_0^{h-j}[\Theta_T], \bar{E}_0^{1+j}[\Theta_T]\right). \tag{38}$$

From the previous condition, for $k \geq 1$, we have[7]

$$\beta_{2k} = \frac{Cov\left(\bar{E}_0^k[\Theta_T], \bar{E}_0^{k+1}[\Theta_T]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)} = \frac{Cov\left(\bar{E}_0^k[\Theta_T], E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)} = \frac{Var\left(E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)} \geq 0 \;\; \forall i, \tag{39}$$

where the second equation follows from symmetry and the last equation follows from the law of iterated expectations. Similarly, for $k \geq 1$, we have

$$\beta_{2k-1} = \frac{Cov\left(\bar{E}_0^k[\Theta_T], \bar{E}_0^k[\Theta_T]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)} = \frac{Var\left(\bar{E}_0^k[\Theta_T]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)} \geq 0. \tag{40}$$

Now, note that for any random variable X, and any information I, according to the law of total variance, we have:

$$Var\left(\mathbb{E}[X|I]\right) \leq Var\left(X\right).$$

As a result, $Var\left(\bar{E}_0^{k+1}[\Theta_T]\right) = Var\left(E\left[E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]|s\right]\right) \leq Var\left(E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right)$ and $Var\left(E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right) = Var\left(E\left[\bar{E}_0^k[\Theta_T]|\omega_i\right]\right) \leq Var\left(\bar{E}_0^k[\Theta_T]\right)$, where, as a reminder, s is the aggregate state of the Nature and ω_i is the information of agent i. Together with conditions (39) and (40), we know that, for all $k \geq 1$, $\beta_{2k+1} \leq \beta_{2k} \leq \beta_{2k-1}$. This proves that, for all $h \geq 2$, $\beta_h \in [0, 1]$ and is weakly decreasing in h.

Now we try to prove β_h is strictly decreasing in h. Note that from condition (39), $\beta_2 = \frac{Var\left(E_{i,0}\left[\bar{E}_0[\Theta_T]\right]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)} \leq \beta_1 = 1$. If $\beta_2 = \beta_1 = 1$, we have $Var\left(E_{i,0}\left[\bar{E}_0[\Theta_T]\right]\right) = Var\left(\bar{E}_0[\Theta_T]\right)$ for all i. This means that

$$\begin{aligned} Var\left(E_{i,0}\left[\bar{E}_0[\Theta_T]\right] - \bar{E}_0[\Theta_T]\right) &= Var\left(E_{i,0}\left[\bar{E}_0[\Theta_T]\right]\right) + Var\left(\bar{E}_0[\Theta_T]\right) - 2Cov\left(E_{i,0}\left[\bar{E}_0[\Theta_T]\right], \bar{E}_0[\Theta_T]\right) \\ &= Var\left(\bar{E}_0[\Theta_T]\right) - Var\left(E_{i,0}\left[\bar{E}_0[\Theta_T]\right]\right) = 0, \end{aligned} \tag{41}$$

[7] Note that under incomplete information, we have $Var\left(\bar{E}_0^1[\Theta_T]\right) > 0$, so all β_k is well defined. To prove it, note that if $Var\left(\bar{E}_0^1[\Theta_T]\right) = 0$, together with the fact that the mean of Θ_T is zero, we have $\bar{E}_0^1[\Theta_T] = 0$ almost surely. As a result, we have $Var\left(E_{i,0}\left[\Theta_T\right]\right) = Cov\left(\Theta_T, E_{i,0}\left[\Theta_T\right]\right) = Cov\left(\Theta_T, \bar{E}_0\left[\Theta_T\right]\right) = 0$, and $E_{i,0}\left[\Theta_T\right] = 0 = E_{j,0}\left[\Theta_T\right]$ almost surely for all i, j. This is inconsistent with the definition about incomplete information in Definition 1.

where the second equality follows from the law of iterated expectations. As a result, for all i, $E_{i,0}\left[\bar{E}_0[\Theta_T]\right] = \bar{E}_0[\Theta_T]$ almost surely. We henceforth have that

$$
\begin{aligned}
Var\left(\bar{E}_0\left[\Theta_T\right]\right) &= Cov\left(\Theta_T, \bar{E}_0^2\left[\Theta_T\right]\right) = Cov\left(\Theta_T, E_{i,0}\left[\bar{E}_0\left[\Theta_T\right]\right]\right) \\
&= Cov\left(\Theta_T, \bar{E}_0\left[\Theta_T\right]\right) = Cov\left(\Theta_T, E_{i,0}\left[\Theta_T\right]\right) = Var\left(E_{i,0}\left[\Theta_T\right]\right),
\end{aligned}
$$

where the first equality follows a similar argument as condition (38), the second and fourth equalities follow from symmetry, the third equality follows from $E_{i,0}\left[\bar{E}_0[\Theta_T]\right] = \bar{E}_0[\Theta_T]$ almost surely, and the last equality follows from the law of iterated expectations. This means that

$$
\begin{aligned}
Var\left(E_{i,0}\left[\Theta_T\right] - \bar{E}_0\left[\Theta_T\right]\right) &= Var\left(E_{i,0}\left[\Theta_T\right]\right) + Var\left(\bar{E}_0\left[\Theta_T\right]\right) - Cov\left(E_{i,0}\left[\Theta_T\right], \bar{E}_0\left[\Theta_T\right]\right) \\
&= Var\left(E_{i,0}\left[\Theta_T\right]\right) - Var\left(\bar{E}_0\left[\Theta_T\right]\right) = 0,
\end{aligned} \tag{42}
$$

where the second equality follows from symmetry. As a result, $E_{i,0}\left[\Theta_T\right] = \bar{E}_0\left[\Theta_T\right]$ almost surely for all i, and $E_{i,0}\left[\Theta_T\right] = E_{j,0}\left[\Theta_T\right]$ almost surely for all i, j. This is contradictory to the definition of incomplete information. As a result, $\beta_2 < \beta_1 = 1$.

Now, suppose it is not the case that β_h is strictly decreasing in h. Then there exists a smallest $h^* > 1$ such that $\beta_{h^*+1} = \beta_{h^*}$.

If $h^* = 2k$ for some $k \geq 1$. From conditions (39) and (40), we have $Var\left(E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right) = Var\left(\bar{E}_0^{k+1}[\Theta_T]\right)$. Following a similar argument as condition (42), we have $E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right] = \bar{E}_0^{k+1}[\Theta_T]$ almost surely. We henceforth have

$$
\begin{aligned}
Cov\left(\bar{E}_0^k[\Theta_T], \bar{E}_0^{k+1}[\Theta_T]\right) &= Cov\left(E_{i,0}\left[\bar{E}_0^{k-1}[\Theta_T]\right], \bar{E}_0^{k+1}[\Theta_T]\right) = Cov\left(E_{i,0}\left[\bar{E}_0^{k-1}[\Theta_T]\right], E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right) \\
&= Cov\left(E_{i,0}\left[\bar{E}_0^{k-1}[\Theta_T]\right], \bar{E}_0^k[\Theta_T]\right) = Cov\left(\bar{E}_0^k[\Theta_T], \bar{E}_0^k[\Theta_T]\right),
\end{aligned}
$$

where the first and the last equalities follow from symmetry, the second equality follows from the fact that $E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right] = \bar{E}_0^{k+1}[\Theta_T]$ almost surely, and the third equality follows from the law of iterated expectations. This expression means $\beta_{h^*-1} = \beta_{2k-1} = \beta_{2k} = \beta_{h^*}$, which contradicts the fact that h^* is the smallest h such that $\beta_{h^*+1} = \beta_{h^*}$.

If $h^* = 2k-1$ for some $k \geq 2$, from conditions (39) and (40), we have $Var\left(E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right]\right) = Var\left(\bar{E}_0^k[\Theta_T]\right)$. Following a similar argument as condition (41) for all i, $E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right] = \bar{E}_0^k[\Theta_T]$ almost surely. We henceforth have

$$
\begin{aligned}
Var\left(\bar{E}_0^k[\Theta_T]\right) &= Cov\left(\bar{E}_0^{k-1}\left[\Theta_T\right], \bar{E}_0^{k+1}\left[\Theta_T\right]\right) = Cov\left(\bar{E}_0^{k-1}\left[\Theta_T\right], E_{i,0}\left[\bar{E}_0^k\left[\Theta_T\right]\right]\right) \\
&= Cov\left(\bar{E}_0^{k-1}\left[\Theta_T\right], \bar{E}_0^k[\Theta_T]\right) = Cov\left(\bar{E}_0^{k-1}\left[\Theta_T\right], E_{i,0}\left[\bar{E}_0^{k-1}\left[\Theta_T\right]\right]\right) \\
&= Var\left(E_{i,0}\left[\bar{E}_0^{k-1}[\Theta_T]\right]\right),
\end{aligned}
$$

where the first equality follows a similar argument as condition (38), the second and forth equalities follow from symmetry, the third equality follows from $E_{i,0}\left[\bar{E}_0^k[\Theta_T]\right] = \bar{E}_0^k[\Theta_T]$ almost surely, and the last equality follows from the law of iterated expectations. This expression means that $\beta_{h^*-1} = \beta_{2k-2} = \beta_{2k-1} = \beta_{h^*}$, which contradicts the fact that h^* is the smallest h such that $\beta_{h^*+1} = \beta_{h^*}$.

As a result, β_h is strictly decreasing in h. This implies that $\beta_h < \beta_1 = 1$, $\forall h \geq 2$. It also means that $\beta_h > 0\ \forall h$. If not, there exists a h^* such that $\beta_{h^*} = 0$. From strict monotonicity, we then have $\beta_{h^*+1} < \beta_{h^*} = 0$, which contradicts $\beta_{h^*+1} \geq 0$. This finishes the proof of Proposition 3.

Proof of Corollary 1. Corollary 1 follows directly from part (ii) of Theorem 1.

Proof of Theorem 2. To simplify notation, we extend the definition of $s_{h,\tau} = \sum_{r=1}^{h} \chi_{r,\tau}$ for all $h > \tau$. In the case that $h > \tau$, from Lemma 2, we have $\chi_{h,\tau} = 0$. As a result, $s_{h,\tau} = s_{\tau,\tau}$ for all $h > \tau$. We also define $s_{0,\tau} = 0$ for all $\tau \geq 1$.
From condition (16), we have

$$s_{h,\tau} = \gamma^{\tau-1}(\gamma+\alpha) + \sum_{l=1}^{\tau-1} \alpha\gamma^{l-1} s_{h-1,\tau-l} \quad \forall h \geq 1 \textit{ and } \tau \geq 1. \tag{43}$$

Now, for all $\tau \geq 1$, as $\chi_{h,\tau} = 0$ for $h > \tau$, we can use $d_\tau = s_{\tau,\tau}$ denote the combined effect of beliefs of all different orders. From condition (43), we have

$$d_\tau = \gamma^{\tau-1}(\gamma+\alpha) + \sum_{l=1}^{\tau-1} \alpha\gamma^{l-1} d_{\tau-l} \quad \forall \tau \geq 1, \tag{44}$$

where we use the fact that $s_{h,\tau} = s_{\tau,\tau}$ for all $h > \tau$. From condition (44), we can easily verify, by induction, that

$$d_\tau = (\gamma+\alpha)^\tau \quad \forall \tau \geq 1. \tag{45}$$

For any $h \geq 1$, we now prove that $s_{h,\tau}/s_{\tau,\tau} = s_{h,\tau}/d_\tau$ strictly decreases with $\tau \geq h$. Notice that from condition (43), we have, for all $\tau \geq h \geq 1$,

$$\begin{aligned} s_{h,\tau+1} &= \gamma^\tau(\gamma+\alpha) + \alpha s_{h-1,\tau} + \sum_{l=1}^{\tau-1} \alpha\gamma^l s_{h-1,\tau-l} \\ &= \gamma s_{h,\tau} + \alpha s_{h-1,\tau} < (\gamma+\alpha) s_{h,\tau}. \end{aligned} \tag{46}$$

Also note that from condition (45), we have $s_{\tau+1,\tau+1} = d_{\tau+1} = (\gamma+\alpha) d_\tau = (\gamma+\alpha) s_{\tau,\tau}$. Together, we have $s_{h,\tau+1}/s_{\tau+1,\tau+1} < s_{h,\tau}/s_{\tau,\tau}$ for all $\tau \geq h \geq 1$.

Finally, we prove that, for any $h \geq 1$, $s_{h,\tau}/s_{\tau,\tau} \to 0$ as $\tau \to +\infty$. Because $s_{1,\tau} = \gamma^{\tau-1}(\gamma+\alpha)$ and $s_{\tau,\tau} = (\gamma+\alpha)^\tau$, $\lim_{\tau\to\infty} s_{h,\tau}/s_{\tau,\tau} \to 0$ holds for $h = 1$. Suppose there is some h such that $\lim_{\tau\to\infty} s_{h,\tau}/s_{\tau,\tau} \to 0$ does not hold, let $h^* > 1$ be the smallest of such h. As $s_{h^*,\tau}/s_{\tau,\tau}$ is strictly decreasing in τ, there exists $\Gamma > 0$

such that $\lim_{\tau\to\infty} s_{h^*,\tau}/s_{\tau,\tau} \to \Gamma$. From conditions (45) and (46), we have $\frac{s_{h^*,\tau+1}}{s_{\tau+1,\tau+1}} = \frac{\gamma}{\gamma+\alpha}\frac{s_{h^*,\tau}}{s_{\tau,\tau}} + \frac{\alpha}{\gamma+\alpha}\frac{s_{h^*-1,\tau}}{s_{\tau,\tau}}$. Let $\tau \to +\infty$, we have $\Gamma = \frac{\gamma}{\gamma+\alpha}\Gamma$. This cannot be true as $\alpha, \gamma > 0$. As a result, $\lim_{\tau\to\infty} s_{h,\tau}/s_{\tau,\tau} \to 0$ for all $h \geq 1$.

Proof of Proposition 4. By Theorem 1, the ratio $\frac{\phi_T}{\phi_T^*}$ is strictly decreasing in T and bounded in $(0,1)$. It follows that $\frac{\phi_T}{\phi_T^*}$ necessarily converges to some $\varphi \in [0,1)$ as $T \to \infty$. Similarly, by Proposition 3, β_h is strictly decreasing in T and bounded in $(0,1)$. It follows that β_h necessarily converges to some $\underline{\beta} \in [0,1)$ as $T \to \infty$.

We first prove $\varphi = \underline{\beta} \equiv \lim_{h\to\infty} \beta_h$. We note that for, any $\vartheta > 0$, there exists a h^*, such that $\left|\beta_h - \underline{\beta}\right| < \frac{\vartheta}{2}$ for all $h \geq h^*$. From Theorem 2, we can then find $T^* \in \mathbb{N}_+$ such that, for all $T \geq T^*$, $\frac{s_{h^*-1,T}}{s_{T,T}} \leq \frac{\vartheta}{2}$. Together with conditions (18) and (20), we have, for all $T \geq \max\{h^*, T^*\}$,

$$\begin{aligned}
\left|\frac{\phi_T}{\phi_T^*} - \underline{\beta}\right| &= \left|\frac{\sum_{h=1}^{T}\chi_{h,T}\left(\beta_h - \underline{\beta}\right)}{s_{T,T}}\right| = \left|\frac{\sum_{h=1}^{h^*-1}\chi_{h,T}\left(\beta_h - \underline{\beta}\right) + \sum_{h=h^*}^{T}\chi_{h,T}\left(\beta_h - \underline{\beta}\right)}{s_{T,T}}\right| \\
&\leq \frac{\sum_{h=1}^{h^*-1}\chi_{h,T}}{s_{T,T}} + \frac{\sum_{h=h^*}^{T}\chi_{h,T}\frac{\vartheta}{2}}{s_{T,T}} \\
&\leq \frac{\vartheta}{2} + \frac{\vartheta}{2} = \vartheta,
\end{aligned}$$

where the first inequality we use the fact that $\left|\beta_h - \underline{\beta}\right| \leq 1$ and the second inequality uses the fact that $\frac{\sum_{h=h^*}^{T}\chi_{h,T}}{s_{T,T}} \leq \frac{s_{T,T}}{s_{T,T}} = 1$. As a result, $\varphi \equiv lim_{T\to\infty}\frac{\phi_T}{\phi_T^*} = \underline{\beta}$.

Finally, from condition (40), we know $\beta_{2h-1} = \frac{Var\left(\bar{E}_0^h[\Theta_T]\right)}{Var\left(\bar{E}_0^1[\Theta_T]\right)}$. If $\lim_{h\to\infty} Var\left(\bar{E}^h[\Theta_T]\right) = 0$, we have $\lim_{h\to\infty}\beta_{2h-1} = 0$. As β_h is decreasing in h, we also have

$$\underline{\beta} = \lim_{h\to\infty}\beta_h = 0. \tag{47}$$

As a result, $lim_{T\to\infty}\frac{\phi_T}{\phi_T^*} = 0$.

Proof of Proposition 5. Under the assumed information structure, we have for any $h \in \{1, ..., T\}$ and $0 \leq t_1 < t_2 < \cdots < t_h < T$,

$$\bar{E}_{t_1}[\bar{E}_{t_2}[...[\bar{E}_{t_h}[\Theta_T]...]] = \lambda^h z. \tag{48}$$

Now we prove by induction that, for all $t \leq T-1$,

$$a_t = (\gamma+\alpha)\left\{\Pi_{\tau=t+1}^{T-1}(\gamma+\lambda\alpha)\,\bar{E}_t[\Theta_T]\right\}. \tag{49}$$

Since $\Theta_t = 0$ for all $t \neq T$, together with condition (16), we have $a_T = \Theta_T$ and $a_{T-1} = (\gamma+\alpha)\,\bar{E}_{T-1}[\Theta_T]$. As a result, condition (49) holds for $t = T-1$. Now, pick a $t \leq T-2$, assume that the claim holds for all $\tau \in \{t+1, ..., T-1\}$, and let us prove that it also holds for t. Using the claim for all $\tau \in \{t+1, ..., T-1\}$, condition (16), and condition (48), we have, for $t \leq T-2$,

$$a_t = \gamma^{T-t}\bar{E}_t\left[\Theta_T\right] + \alpha\bar{E}_t\left[a_{t+1}\right] + \alpha\sum_{k=2}^{T-t}\gamma^{k-1}\bar{E}_t\left[a_{t+k}\right],$$

$$\gamma\bar{E}_t\left[a_{t+1}\right] = \lambda\gamma^{T-t}\bar{E}_t\left[\Theta_T\right] + \lambda\alpha\sum_{k=2}^{T-t}\gamma^{k-1}\bar{E}_t\left[a_{t+k}\right].$$

As a result, we have $a_t = \left(\frac{\gamma}{\lambda}+\alpha\right)\bar{E}_t\left[a_{t+1}\right]$. Together with condition (49) for $t+1$, we have

$$a_t = (\gamma+\alpha)\left\{\Pi_{\tau=t+1}^{T-1}\left(\gamma+\lambda\alpha\right)\bar{E}_t[\Theta_T]\right\}.$$

This proves condition (49) for all $t \leq T-1$. As a result, $\phi_T = (\gamma+\alpha)\,\Pi_{t=1}^{T-1}\left(\gamma+\lambda\alpha\right)$. This proves Proposition 5.

Proof of Corollary 2. Note that $\delta-\delta' = \alpha\left(1-\lambda\right)$ increases with α for any given $\lambda < 1$.

Proof of Lemma 3. As firms have complete information, the canonical NKPC in condition (10) holds. Substituting it into the consumption beauty contest, condition (5), and using the fact that future markup shocks are unpredictable, we have

$$\tilde{y}_t = -\sigma\tilde{R}_t - \sigma\sum_{k=1}^{\infty}\beta^k\bar{E}_t^c[\tilde{R}_{t+k}] + \sum_{k=1}^{\infty}\left(1-\beta+k\sigma\kappa\right)\beta^{k-1}\bar{E}_t^c[\tilde{y}_{t+k}].$$

Proof of Proposition 6. Let $\left\{\tilde{y}_t^{trap}, \tilde{\pi}_t^{trap}\right\}_{t=0}^{T}$ denote the liquidity-trap level of output and inflation (i.e., the one obtained when the period-T nominal interest rate is fixed at the steady-state value, $\tilde{R}_T = 0$). From conditions (9) and (10), we have, for all $t \leq T-1$,

$$\tilde{y}_t - \tilde{y}_t^{trap} = \sigma E_t[\tilde{\pi}_{t+1} - \tilde{\pi}_{t+1}^{trap}] + E_t[\tilde{y}_{t+1} - \tilde{y}_{t+1}^{trap}], \tag{50}$$

$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} = \kappa\left(\tilde{y}_t - \tilde{y}_t^{trap}\right) + \beta E_t[\tilde{\pi}_{t+1} - \tilde{\pi}_{t+1}^{trap}]. \tag{51}$$

Now we will prove the following stronger result, which nests the representation in condition (24): there exists positive scalars $\{\phi_\tau^*, \varpi_\tau^*\}_{\tau\geq 0}$ such that, whenever Assumptions 2 hold and z is commonly known, the equilibrium spending and inflation at any $t \leq T$ are given by

$$\tilde{y}_t - \tilde{y}_t^{trap} = -\phi_{T-t}^* \cdot E_t[\tilde{R}_T], \tag{52}$$

$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} = \kappa\left(\tilde{y}_t - \tilde{y}_t^{trap}\right) - \varpi_{T-t}^* \cdot E_t[\tilde{R}_T]. \tag{53}$$

We prove this result by induction, starting with $t = T$ and proceeding backwards. When $t = T$, under Assumption 2, we have $\tilde{y}_T - \tilde{y}_T^{trap} = -\sigma\tilde{R}_T$ and $\tilde{\pi}_T - \tilde{\pi}_T^{trap} = \kappa\left(\tilde{y}_T - \tilde{y}_T^{trap}\right)$. This verifies (52) and (53) for

$t = T$, with

$$\phi_0^* = \sigma \text{ and } \varpi_0^* = 0. \tag{54}$$

Now suppose that the result holds for arbitrary $t \in \{1, ..., T\}$ and let's prove that it also holds for $t-1$. By the assumption that (52) and (53) hold at t along with the Law of Iterated Expectations, we have

$$E_{t-1}[\tilde{y}_t - \tilde{y}_t^{trap}] = -\phi_{T-t}^* \cdot E_{t-1}[\tilde{R}_T],$$
$$E_{t-1}[\tilde{\pi}_t - \tilde{\pi}_t^{trap}] = -\left(\kappa\phi_{T-t}^* + \varpi_{T-t}^*\right) \cdot E_{t-1}[\tilde{R}_T].$$

Using the above together with conditions (50) and (51) verifies that (52) and (53) hold also for $t-1$, with

$$\phi_{T-t+1}^* = (1+\sigma\kappa)\,\phi_{T-t}^* + \sigma\varpi_{T-t}^*, \tag{55}$$

$$\varpi_{T-t+1}^* = \beta\kappa\phi_{T-t}^* + \beta\varpi_{T-t}^*. \tag{56}$$

This completes the proof of conditions (50) and (51), and gives a recursive formula that can be used to compute ϕ_T^*.

Now we prove the Proposition. From conditions (55) and (56), we have that, for all $\tau \geq 0$,

$$\phi_{\tau+1}^* = (1+\sigma\kappa)\,\phi_\tau^* + \sigma\varpi_\tau^*, \tag{57}$$

$$\varpi_{\tau+1}^* = \beta\kappa\phi_\tau^* + \beta\varpi_\tau^*. \tag{58}$$

Together with condition (54), we know, as $\kappa > 0$, $\varpi_\tau^* > 0, \ \ \forall\tau \geq 1$. Then, from condition (57), we have $\phi_\tau^* > \sigma, \ \forall\tau \geq 1$, and ϕ_τ^* is strictly increasing in τ. Moreover, as $1+\sigma\kappa > 1$, we know ϕ_τ^* explodes to infinity as $\tau \to \infty$ from condition (57).

Finally, we prove a few more results useful for the rest of the paper. First, we prove a recursive relationship about $\{\phi_\tau^*\}_{\tau\geq 0}$.

$$\frac{\phi_{\tau+1}^*}{\phi_\tau^*} + \beta\frac{\phi_{\tau-1}^*}{\phi_\tau^*} = 1 + \beta + \sigma\kappa \quad \forall\tau \geq 1. \tag{59}$$

From condition (57), we have, for all $\tau \geq 1$,

$$\beta\phi_\tau^* = \beta\,(1+\sigma\kappa)\,\phi_{\tau-1}^* + \sigma\beta\varpi_{\tau-1}^*.$$

Together with conditions (57) and (58), we arrive at condition (59).

Second, we prove that, when $\kappa > 0$,

$$\frac{\phi_\tau^*}{\phi_{\tau-1}^*} \text{ is strictly increasing in } \tau \geq 1. \tag{60}$$

From conditions (57) and (58), we have $\phi_1^* = \sigma\,(1+\sigma\kappa)$ and $\phi_2^* = \sigma\left((1+\sigma\kappa)^2 + \sigma\kappa\beta\right)$. As a result, when

$\kappa > 0$,

$$\frac{\phi_2^*}{\phi_1^*} = 1 + \sigma\kappa + \frac{\sigma\kappa\beta}{1+\sigma\kappa} > \frac{\phi_1^*}{\phi_0^*}.$$

Now we proceed by induction. Suppose that, for $\tau \geq 1$, we have $\frac{\phi_{\tau+1}^*}{\phi_\tau^*} > \frac{\phi_\tau^*}{\phi_{\tau-1}^*}$. Using condition (59) for τ and $\tau+1$, we have $\frac{\phi_{\tau+2}^*}{\phi_{\tau+1}^*} > \frac{\phi_{\tau+1}^*}{\phi_\tau^*}$. This proves (60).

Finally, from condition (59), we know $\frac{\phi_\tau^*}{\phi_{\tau-1}^*}$ is bounded above. Together with (60), $\frac{\phi_\tau^*}{\phi_{\tau-1}^*}$ must converge to $\Gamma^* > 0$, as $\tau \to \infty$. From condition (59) again, we know Γ^* satisfy

$$\Gamma^* + \beta\frac{1}{\Gamma^*} = 1 + \beta + \sigma\kappa. \tag{61}$$

Proof of Proposition 7. With $\left\{\tilde{y}_t^{trap}, \tilde{\pi}_t^{trap}\right\}_{t=0}^T$ defined as in the proof of Proposition 6, along with the fact that it is common knowledge monetary policy replicates flexible-price allocations from $T+1$ and on, we can rewrite the two beauty contests as follows:

$$\tilde{y}_t - \tilde{y}_t^{trap} = -\sigma\beta^{T-t}\bar{E}_t^c[\tilde{R}_T] + \sum_{k=1}^{T-t}\sigma\beta^{k-1}\bar{E}_t^c\left[\tilde{\pi}_{t+k} - \tilde{\pi}_{t+k}^{trap}\right] + (1-\beta)\sum_{k=1}^{T-t}\beta^{k-1}\bar{E}_t^c\left[\tilde{y}_{t+k} - \tilde{y}_{t+k}^{trap}\right], \tag{62}$$

$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} = \kappa\left(\tilde{y}_t - \tilde{y}_t^{trap}\right) + \kappa\sum_{k=1}^{T-t}(\beta\theta)^k\,\bar{E}_t^f\left[\tilde{y}_{t+k} - \tilde{y}_{t+k}^{trap}\right] + \tfrac{1-\theta}{\theta}\sum_{k=1}^{T-t}(\beta\theta)^k\,\bar{E}_t^f\left[\tilde{\pi}_{t+k} - \tilde{\pi}_{t+k}^{trap}\right]. \tag{63}$$

Consider the following claim, which nests the representation in condition (25): under Assumption 3, there exists functions $\phi, \varpi : (0,1] \times (0,1] \times \mathbb{N} \to \mathbb{R}_+$ such that, for any $t \leq T-1$,

$$\begin{aligned}
\tilde{y}_t - \tilde{y}_t^{trap} &= -\phi\left(\lambda_c, \lambda_f, T-t\right)\bar{E}_t^c[\tilde{R}_T], && (64)\\
\tilde{\pi}_t - \tilde{\pi}_t^{trap} &= \kappa\left(\tilde{y}_t - \tilde{y}_t^{trap}\right) - \varpi\left(\lambda_c, \lambda_f, T-t\right)\bar{E}_t^f[\tilde{R}_T]. && (65)
\end{aligned}$$

We now establish this claim by induction.

First, consider $t = T$, as $\tilde{R}_T$ becomes common known at period T, we have

$$\tilde{y}_T - \tilde{y}_T^{trap} = -\sigma\tilde{R}_T \qquad \text{and} \qquad \tilde{\pi}_T - \tilde{\pi}_T^{trap} = \kappa\left(\tilde{y}_T - \tilde{y}_T^{trap}\right).$$

Then, consider $t = T-1$. From conditions (62) and (63), we have

$$\begin{aligned}
&\tilde{y}_{T-1} - \tilde{y}_{T-1}^{trap} = -\sigma(1+\sigma\kappa)\bar{E}_{T-1}^c[\tilde{R}_T],\\
&\tilde{\pi}_{T-1} - \tilde{\pi}_{T-1}^{trap} = \kappa\left(\tilde{y}_{T-1} - \tilde{y}_{T-1}^{trap}\right) - \sigma\kappa\beta\bar{E}_{T-1}^f[\tilde{R}_T].
\end{aligned}$$

It follows that the claim holds for $t = T-1$ with

$$\phi(\lambda_c, \lambda_f, 1) = \sigma\left(1+\sigma\kappa\right) \qquad \text{and} \qquad \varpi(\lambda_c, \lambda_f, 1) = \sigma\kappa\beta. \tag{66}$$

Now, pick an arbitrary $t \leq T-2$, assume that conditions (64) and (65) hold for all $\tau \in \{t+1, ..., T-1\}$, and let us prove that it also holds for t. Since the claim holds for $\tau \in \{t+1, ...T-1\}$, and since $\tilde{y}_T - \tilde{y}_T^{trap} = -\sigma \tilde{R}_T$ and $\tilde{\pi}_T - \tilde{\pi}_T^{trap} = -\kappa\sigma\tilde{R}_T$, from condition (62), we have

$$\begin{aligned}\tilde{y}_t - \tilde{y}_t^{trap} = &-\sigma\beta^{T-t-1}\left(1+\sigma\kappa\right)\bar{E}_t^c[\tilde{R}_T] - \left(1-\beta+\sigma\kappa\right)\sum_{k=1}^{T-t-1}\beta^{k-1}\phi\left(\lambda_c, \lambda_f, T-t-k\right)\bar{E}_t^c[\bar{E}_{t+k}^c[\tilde{R}_T]] \\ &-\sigma\sum_{k=1}^{T-t-1}\beta^{k-1}\varpi\left(\lambda_c, \lambda_f, T-t-k\right)\bar{E}_t^c\left[\bar{E}_{t+k}^f[\tilde{R}_T]\right].\end{aligned}$$

As a result, we have

$$\begin{aligned}\tilde{y}_t - \tilde{y}_t^{trap} = &-\sigma\beta^{T-t-1}\left(1+\sigma\kappa\right)\bar{E}_t^c[\tilde{R}_T] \\ &-\sum_{k=1}^{T-t-1}\beta^{k-1}\left[\left(1-\beta+\sigma\kappa\right)\lambda_c\phi\left(\lambda_c, \lambda_f, T-t-k\right) + \sigma\lambda_f\varpi\left(\lambda_c, \lambda_f, T-t-k\right)\right]\bar{E}_t^c[\tilde{R}_T],\end{aligned}$$

where we have used the fact that, under Assumption 3, for $1 \leq k \leq T-t-1$,

$$\bar{E}_t^c[\bar{E}_{t+k}^c[\tilde{R}_T]] = \lambda_c\bar{E}_t^c[\tilde{R}_T] \quad \text{and} \quad \bar{E}_t^c[\bar{E}_{t+k}^f[\tilde{R}_T]] = \lambda_f\bar{E}_t^c[\tilde{R}_T]. \tag{67}$$

This proves the part of the claim that regards output, condition (64), with

$$\phi\left(\lambda_c, \lambda_f, T-t\right) = \beta^{T-t-1}\left(\sigma+\sigma^2\kappa\right) + \sum_{k=1}^{T-t-1}\beta^{k-1}\left[\left(1-\beta+\sigma\kappa\right)\lambda_c\phi\left(\lambda_c, \lambda_f, T-t-k\right) + \sigma\lambda_f\varpi\left(\lambda_c, \lambda_f, T-t-k\right)\right]. \tag{68}$$

Similarly, the inflation beauty contest in condition (63) gives

$$\begin{aligned}\tilde{\pi}_t - \tilde{\pi}_t^{trap} = &\kappa\left(\tilde{y}_t - \tilde{y}_t^{trap}\right) - \sigma\tfrac{\kappa}{\theta}\left(\beta\theta\right)^{T-t}\bar{E}_t^f[\tilde{R}_T] - \tfrac{\kappa}{\theta}\sum_{k=1}^{T-t-1}\left(\beta\theta\right)^k\phi\left(\lambda_c, \lambda_f, T-t-k\right)\bar{E}_t^f[\bar{E}_{t+k}^c[\tilde{R}_T]] \\ &-\tfrac{1-\theta}{\theta}\sum_{k=1}^{T-t-1}\left(\beta\theta\right)^k\varpi\left(\lambda_c, \lambda_f, T-t-k\right)\bar{E}_t^f[\bar{E}_{t+k}^f[\tilde{R}_T]].\end{aligned}$$

As a result, we have

$$\begin{aligned}\tilde{\pi}_t - \tilde{\pi}_t^{trap} = &\kappa\left(\tilde{y}_t - \tilde{y}_t^{trap}\right) \\ &-\left\{\sigma\tfrac{\kappa}{\theta}\left(\beta\theta\right)^{T-t} + \sum_{k=1}^{T-t-1}\left(\beta\theta\right)^k\left[\tfrac{\kappa\lambda_c}{\theta}\phi\left(\lambda_c, \lambda_f, T-t-k\right) + \tfrac{(1-\theta)\lambda_f}{\theta}\varpi\left(\lambda_c, \lambda_f, T-t-k\right)\right]\right\}\bar{E}_t^f[\tilde{R}_T],\end{aligned}$$

where we have used the fact that, similarly to the consumers' case, for $1 \leq k \leq T-t-1$,

$$\bar{E}_t^f[\bar{E}_{t+k}^c[\tilde{R}_T]] = \lambda_c \bar{E}_t^f[\tilde{R}_T] \quad \text{and} \quad \bar{E}_t^f[\bar{E}_{t+k}^f[\tilde{R}_T]] = \lambda_f \bar{E}_t^f[\tilde{R}_T]. \tag{69}$$

This proves the part of the claim that regards inflation, condition (65) with

$$\varpi\left(\lambda_c, \lambda_f, T-t\right) = \sigma\frac{\kappa}{\theta}\left(\beta\theta\right)^{T-t} + \sum_{k=1}^{T-t-1}\left(\beta\theta\right)^k\left(\tfrac{\kappa\lambda_c}{\theta}\phi\left(\lambda_c, \lambda_f, T-t-k\right) + \tfrac{(1-\theta)\lambda_f}{\theta}\varpi\left(\lambda_c, \lambda_f, T-t-k\right)\right). \tag{70}$$

We finally provide a recursive formula for computing $\phi(\lambda_c, \lambda_f, T-t)$ and $\varpi\left(\lambda_c, \lambda_f, T-t\right)$, which will be useful later. From condition (68), we have, for $t \leq T-2$,

$$\begin{aligned}\phi\left(\lambda_c, \lambda_f, T-t\right) &= \beta\phi\left(\lambda_c, \lambda_f, T-t-1\right) + \left(1-\beta+\sigma\kappa\right)\lambda_c\phi\left(\lambda_c, \lambda_f, T-t-1\right) + \sigma\lambda_f\varpi\left(\lambda_c, \lambda_f, T-t-1\right) \\ &= \left(\beta + \left(1-\beta+\sigma\kappa\right)\lambda_c\right)\phi\left(\lambda_c, \lambda_f, T-t-1\right) + \sigma\lambda_f\varpi\left(\lambda_c, \lambda_f, T-t-1\right). \end{aligned} \tag{71}$$

Similarly, from condition (70), we have, for $t \leq T-2$,

$$\begin{aligned}\varpi\left(\lambda_c, \lambda_f, T-t\right) &= \beta\theta\varpi\left(\lambda_c, \lambda_f, T-t-1\right) + \beta\theta\left(\tfrac{\kappa\lambda_c}{\theta}\phi\left(\lambda_c, \lambda_f, T-t-1\right) + \tfrac{(1-\theta)\lambda_f}{\theta}\varpi\left(\lambda_c, \lambda_f, T-t-1\right)\right) \\ &= \kappa\beta\lambda_c\phi\left(\lambda_c, \lambda_f, T-t-1\right) + \beta\left[\theta + \left(1-\theta\right)\lambda_f\right]\varpi\left(\lambda_c, \lambda_f, T-t-1\right). \end{aligned} \tag{72}$$

From now on, to simplify notation, we use ϕ_τ and ϖ_τ as shortcuts for, respectively, $\phi\left(\lambda_c, \lambda_f, \tau\right)$ and $\varpi\left(\lambda_c, \lambda_f, \tau\right)$.

We first prove part (i) of Proposition 7. From condition (68), we know $\phi_\tau > \sigma\beta^\tau$. The fact that $\phi_\tau < \phi_\tau^*$ is a direct corollary from the monotonicity of ϕ_τ with respect to λ_c and λ_f, which will be proved shorty.

We then prove part (ii) of Proposition 7. As $\kappa > 0$, from conditions (66), (71) and (72), we know that $\phi_\tau, \varpi_\tau > 0$ for all $\tau \geq 1$.

We will first prove, for $\tau \geq 2$, $\phi_\tau = \phi\left(\lambda_c, \lambda_f, \tau\right)$ is strictly increasing in both λ_c and λ_f. We will proceed by induction on τ. For $\tau = 2$, from (66), (71) and (72), we have ϕ_2 and ϖ_2 is strictly increasing in both λ_c and λ_f. Suppose for $\tau \geq 2$, ϕ_τ, ϖ_τ is strictly increasing in both λ_c and λ_f. From conditions (71) and (72), we know $\phi_{\tau+1}$ and $\varpi_{\tau+1}$ are strictly increasing in both λ_c and λ_f, where we use the fact that $\phi_\tau, \varpi_\tau > 0$. This proves that, for $\tau \geq 2$, $\phi_\tau = \phi\left(\lambda_c, \lambda_f, \tau\right)$ is strictly increasing in both λ_c and λ_f. Because of the strict monotonicity, we have, for $\tau \geq 2$, whenever $\lambda_c < 1$ and/or $\lambda_f < 1$, $\frac{\phi_\tau}{\phi_\tau^*} = \frac{\phi(\lambda_c, \lambda_f, \tau)}{\phi(1,1,\tau)} < 1$.

We now prove that, whenever $\lambda_c < 1$ and/or $\lambda_f < 1$, the ratio $\frac{\phi_\tau}{\phi_\tau^*} = \frac{\phi(\lambda_c, \lambda_f, \tau)}{\phi_\tau^*}$ is strictly decreasing in $\tau \geq 1$. We start by noticing, from the proof of Proposition 6, we have, for $\tau \geq 3$,

$$\frac{\phi_\tau^*}{\phi_{\tau-1}^*} + \beta\frac{\phi_{\tau-2}^*}{\phi_{\tau-1}^*} = 1 + \beta + \sigma\kappa \quad . \tag{73}$$

Now we prove that ϕ_τ satisfies an inequality with a similar form as (73):

$$\frac{\phi_\tau}{\phi_{\tau-1}} + \beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \leq 1 + \beta + \sigma\kappa\lambda_c \leq 1 + \beta + \sigma\kappa \qquad \forall\tau \geq 3. \tag{74}$$

From condition (71), we have, for $\tau \geq 3$,

$$\begin{aligned} \phi_\tau &= (\beta + (1-\beta)\lambda_c)\,\phi_{\tau-1} + \sigma\kappa\lambda_c\phi_{\tau-1} + \sigma\lambda_f\varpi_{\tau-1}, \\ \tfrac{\beta}{\beta+(1-\beta)\lambda_c}\phi_{\tau-1} &= \beta\phi_{\tau-2} + \tfrac{\sigma\beta\kappa\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{\tau-2} + \tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{\tau-2}. \end{aligned}$$

From the previous two conditions, we have, for $\tau \geq 3$,

$$\begin{aligned} \phi_\tau + \beta\phi_{\tau-2} = (\beta + (1-\beta)\lambda_c)\,\phi_{\tau-1} + \sigma\kappa\lambda_c\phi_{\tau-1} + \sigma\lambda_f\varpi_{\tau-1} \\ + \tfrac{\beta}{\beta+(1-\beta)\lambda_c}\phi_{\tau-1} - \tfrac{\sigma\beta\kappa\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{\tau-2} - \tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{\tau-2}. \end{aligned} \tag{75}$$

Note that, for $\tau \geq 3$ and $\lambda_c, \lambda_f \in (0,1]$, we have

$$\left[(\beta + (1-\beta)\lambda_c) + \sigma\kappa\lambda_c + \tfrac{\beta}{\beta+(1-\beta)\lambda_c}\right]\phi_{\tau-1} \leq (1 + \beta + \sigma\kappa\lambda_c)\,\phi_{\tau-1},$$

and, from condition (72),

$$\begin{aligned} &\sigma\lambda_f\varpi_{\tau-1} - \tfrac{\sigma\beta\kappa\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{\tau-2} - \tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{\tau-2} \\ =&\sigma\lambda_f\left(\kappa\beta\lambda_c\phi_{\tau-2} + \beta\left[\theta + (1-\theta)\lambda_f\right]\varpi_{\tau-2}\right) - \tfrac{\sigma\beta\kappa\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{\tau-2} - \tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{\tau-2} \\ =&\sigma\kappa\beta\lambda_c\left(\lambda_f - \tfrac{1}{\beta+(1-\beta)\lambda_c}\right)\phi_{\tau-2} + \sigma\beta\lambda_f\left[\theta + (1-\theta)\lambda_f - \tfrac{1}{\beta+(1-\beta)\lambda_c}\right]\varpi_{\tau-2} \\ \leq&0. \end{aligned}$$

Together with condition (75), we arrive at condition (74).

Now we can prove that, whenever $\lambda_c < 1$ and/or $\lambda_f < 1$, $\frac{\phi_\tau}{\phi_\tau^*}$ is strictly decreasing in τ. We already prove $\frac{\phi_2}{\phi_2^*} < 1 = \frac{\phi_1}{\phi_1^*}$. We proceed by induction on τ. If $\frac{\phi_\tau}{\phi_\tau^*} < \frac{\phi_{\tau-1}}{\phi_{\tau-1}^*}$ for $\tau \geq 2$, we have $\frac{\phi_{\tau-1}}{\phi_\tau} > \frac{\phi_{\tau-1}^*}{\phi_\tau^*}$. From (73) and (74), we have $\frac{\phi_{\tau+1}}{\phi_\tau} < \frac{\phi_{\tau+1}^*}{\phi_\tau^*}$ and thus $\frac{\phi_{\tau+1}}{\phi_{\tau+1}^*} < \frac{\phi_\tau}{\phi_\tau^*}$. This finishes the proof that $\frac{\phi_\tau}{\phi_\tau^*}$ is strictly decreasing in $\tau \geq 1$, whenever $\lambda_c < 1$ and/or $\lambda_f < 1$.

Now we prove that, whenever $\lambda_c < 1$ and/or $\lambda_f < 1$, $\frac{\phi_\tau}{\phi_\tau^*}$ converges to 0 as $\tau \to \infty$. Because $\frac{\phi_\tau}{\phi_\tau^*} > 0$ is strictly decreasing in $\tau \geq 1$, there exists $\Gamma \in [0,1)$ such that $\frac{\phi_\tau}{\phi_\tau^*} \to \Gamma$ as $\tau \to \infty$. We next prove by contradiction that $\Gamma = 0$.

Suppose first that $\lambda_c < 1$. If $\Gamma > 0$, we have $\frac{\phi_\tau}{\phi_\tau^*}\frac{\phi_{\tau-1}^*}{\phi_{\tau-1}} \to 1$ as $\tau \to \infty$. Because $\frac{\phi_\tau^*}{\phi_{\tau-1}^*} \to \Gamma^*$, we have $\frac{\phi_\tau}{\phi_{\tau-1}} \to \Gamma^*$ and $\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to \frac{1}{\Gamma^*}$ as $\tau \to \infty$. From condition (61), we have $\frac{\phi_\tau}{\phi_{\tau-1}} + \beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to 1 + \beta + \sigma\kappa$ as $\tau \to \infty$. However, this is inconsistent with (74) when $\lambda_c < 1$. As a result, $\Gamma = 0$.

Suppose next that $\lambda_c = 1$ but $\lambda_f < 1$. We prove a stronger version of (74):

$$\frac{\phi_\tau}{\phi_{\tau-1}} + (1 + \sigma\kappa(1-\lambda_f))\,\beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \leq 1 + \beta + \sigma\kappa \quad \forall \tau \geq 3. \tag{76}$$

From conditions (71) and (72), we have, for $\tau \geq 3$,

$$\begin{aligned}
\phi_\tau &= (1+\sigma\kappa)\,\phi_{\tau-1} + \sigma\lambda_f\varpi_{\tau-1},\\
\beta\phi_{\tau-1} &= \beta\phi_{\tau-2} + \beta\sigma\kappa\phi_{\tau-2} + \beta\sigma\lambda_f\varpi_{\tau-2},\\
\varpi_{\tau-1} &= \kappa\beta\phi_{\tau-2} + \beta\,[\theta + (1-\theta)\,\lambda_f]\,\varpi_{\tau-2}.
\end{aligned}$$

As a result, for $\tau \geq 3$,

$$\begin{aligned}
\phi_\tau + \beta\phi_{\tau-2} &= (1+\sigma\kappa+\beta)\,\phi_{\tau-1} + \sigma\lambda_f\varpi_{\tau-1} - \beta\sigma\kappa\phi_{\tau-2} - \beta\sigma\lambda_f\varpi_{\tau-2}\\
&\leq (1+\sigma\kappa+\beta)\,\phi_{\tau-1} + \sigma\,(\lambda_f - 1)\,\kappa\beta\phi_{\tau-2}.
\end{aligned}$$

This proves (76).

Now, if $\Gamma > 0$, similarly, we have $\frac{\phi_\tau}{\phi^*_\tau}\frac{\phi^*_{\tau-1}}{\phi_{\tau-1}} \to 1$ as $\tau \to \infty$. Because $\frac{\phi^*_\tau}{\phi^*_{\tau-1}} \to \Gamma^*$, we have $\frac{\phi_\tau}{\phi_{\tau-1}} \to \Gamma^*$ and $\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to \frac{1}{\Gamma^*}$ as $\tau \to \infty$. From condition (61), we have $\frac{\phi_\tau}{\phi_{\tau-1}} + (1+\sigma\kappa(1-\lambda_f))\,\beta\frac{\phi_{\tau-2}}{\phi_{\tau-1}} \to 1 + \beta + \sigma\kappa + \sigma\kappa\,(1-\lambda_f)\,\beta\frac{1}{\Gamma^*}$ as $\tau \to \infty$. However, this is inconsistent with equation (76) when $\lambda_f < 1$. As a result, $\Gamma = 0$ when $\lambda_c = 1$, but $\lambda_f < 1$.

Finally, we prove that, when λ_c is sufficiently low, $\phi(\lambda_c, \lambda_f, \tau)$ converges to zero as $\tau \to \infty$. The eigenvalues of the dynamic system (ϕ_τ, ϖ_τ) based on conditions (71) and (72) are

$$m_1 = \frac{\beta + (1-\beta+\sigma\kappa)\,\lambda_c + \beta\,[(1-\theta)\,\lambda_f + \theta] - \sqrt{(\beta + (1-\beta+\sigma\kappa)\,\lambda_c - \beta\,[(1-\theta)\,\lambda_f + \theta])^2 + 4\sigma\beta\lambda_f\lambda_c\kappa}}{2} > 0;$$

$$m_2 = \frac{\beta + (1-\beta+\sigma\kappa)\,\lambda_c + \beta\,[(1-\theta)\,\lambda_f + \theta] + \sqrt{(\beta + (1-\beta+\sigma\kappa)\,\lambda_c - \beta\,[(1-\theta)\,\lambda_f + \theta])^2 + 4\sigma\beta\lambda_f\lambda_c\kappa}}{2} > m_1.$$

Note that $\lim_{\lambda_c \to 0} m_2 = \beta < 1$. As a result, when λ_c is sufficiently low, both eigenvalues are below 1, which means that $\phi(\lambda_c, \lambda_f, \tau)$ converges to zero as $\tau \to \infty$.

Proof of Proposition 8. We use $\tilde{g}_t$ to denote the amount of government spending at period t. As mentioned in main text, the government spending $\tilde{g}_t$ is financed by lump sum tax at period t, $\tilde{t}_t = \tilde{g}_t$. Similar to the analysis for monetary policy, we assume $\tilde{g}_t$ becomes commonly known at period t and only allow higher-order uncertainty about future $\tilde{g}$.

Similar to the main text, now we start to work with log-linearized representation. Because the introduction of lump-sum tax, the individual budget constraint becomes

$$\sum_{k=0}^{+\infty}\beta^k\tilde{c}_{i,t+k}=\tilde{a}_{i,t}+\sum_{k=0}^{+\infty}\beta^k\left\{\Omega_1\left(\tilde{w}_{i,t+k}+\tilde{n}_{i,t+k}\right)+\Omega_2\tilde{e}_{i,t+k}-(\Omega_1+\Omega_2-1)\tilde{t}_t\right\},$$

where Ω_1 is the ratio of labor income to total income (net of tax) in steady state, Ω_2 is the ratio of dividend income to total income (net of tax) in steady state, and $\Omega_1+\Omega_2-1$ is the ratio of lump sum tax to total income (net of tax) in steady state. On the other hand, the individual optimal labor supply and Euler equation, conditions (28) and (29), still hold here. Together, this gives rise to the optimal expenditure of consumer $i\in\mathcal{I}_c$ in period t,

$$\begin{aligned}\tilde{c}_{i,t}=\tfrac{(1-\beta)\epsilon\sigma}{\epsilon\sigma+\Omega_1}\tilde{a}_{i,t}-\sigma\sum_{k=1}^{+\infty}\beta^kE_{i,t}\left[\tilde{r}_{t+k}\right]+(1-\beta)\left[\tfrac{(\epsilon+1)\sigma\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{w}_{i,t}+\tfrac{\epsilon\sigma\Omega_2}{\epsilon\sigma+\Omega_1}\tilde{e}_{i,t}-\tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{t}_t\right] \\ +(1-\beta)\sum_{k=1}^{+\infty}\beta^kE_{i,t}\left[\tfrac{(\epsilon+1)\sigma\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{w}_{i,t+k}+\tfrac{\epsilon\sigma\Omega_2}{\epsilon\sigma+\Omega_1}\tilde{e}_{i,t+k}-\tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{t}_{t+k}\right].\end{aligned}\tag{77}$$

Using the fact that assets average to zero and that future idiosyncratic shocks are unpredictable, we obtain the following condition for aggregate spending:

$$\begin{aligned}\tilde{c}_t=-\sigma\sum_{k=1}^{+\infty}\beta^k\bar{E}_t^c\left[\tilde{r}_{t+k}\right]+(1-\beta)\left[\tfrac{(\epsilon+1)\sigma\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{w}_t+\tfrac{\epsilon\sigma\Omega_2}{\epsilon\sigma+\Omega_1}\tilde{e}_t-\tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{t}_t\right] \\ +(1-\beta)\sum_{k=1}^{+\infty}\beta^k\bar{E}_t^c\left[\tfrac{(\epsilon+1)\sigma\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{w}_{t+k}+\tfrac{\epsilon\sigma\Omega_2}{\epsilon\sigma+\Omega_1}\tilde{e}_{t+k}-\tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{t}_{t+k}\right].\end{aligned}\tag{78}$$

The firm side, on the other hand, is essentially same as the case without government spending, as a result, condition (6) still holds, but the formula for marginal cost are different. In particular, from the production function (3) and the optimal labor supply condition (28), we have

$$\tilde{mc}_t=\tilde{w}_t=\epsilon\int_{\mathcal{I}_c}\tilde{n}_{i,t}di+\frac{1}{\sigma}\tilde{c}_t=\epsilon\tilde{y}_t+\frac{1}{\sigma}\tilde{c}_t=\left(\epsilon\Omega_3+\frac{1}{\sigma}\right)\tilde{c}_t+\epsilon\left(1-\Omega_3\right)\tilde{g}_t,\tag{79}$$

where $\tilde{y}_t=\Omega_3\tilde{c}_t+\left(1-\Omega_3\right)\tilde{g}_t$, $\Omega_3=\frac{1}{\Omega_1+\Omega_2}$ is the steady state consumption to output ratio, and $1-\Omega_3=\frac{\Omega_1+\Omega_2-1}{\Omega_1+\Omega_2}$ is the steady state government spending to output ratio.[8] As a result, the inflation beauty contest in condition (6) can be written as

$$\tilde{\pi}_t=\kappa\left(\Omega_c\tilde{c}_t+\left(1-\Omega_c\right)\tilde{g}_t\right)+\kappa\sum_{k=1}^{+\infty}\left(\beta\theta\right)^k\bar{E}_t^f\left[\Omega_c\tilde{c}_{t+k}+\left(1-\Omega_c\right)\tilde{g}_{t+k}\right]+\tfrac{1-\theta}{\theta}\sum_{k=1}^{+\infty}\left(\beta\theta\right)^k\bar{E}_t^f\left[\tilde{\pi}_{t+k}\right]+\varkappa\tilde{\mu}_t,\tag{80}$$

where $\Omega_c=\frac{\epsilon\Omega_3+\frac{1}{\sigma}}{\epsilon+\frac{1}{\sigma}}$.

[8] In steady state, the ratio of government spending to consumption will be equal to ratio of lump sum tax to total income (net of tax), $\Omega_1+\Omega_2-1$. This explains the formula for Ω_3.

Finally, note that the real profit of monopolist j at period t is given by $e_t^j = \left(\frac{p_t^j}{p_t} - w_t^j\right) y_t^j$. After log-linearization, we have $\tilde{e}_t = -\frac{\frac{\Omega_1}{\Omega_1+\Omega_2}}{1-\frac{\Omega_1}{\Omega_1+\Omega_2}}\tilde{w}_t + \tilde{y}_t = -\frac{\Omega_1}{\Omega_2}\tilde{w}_t + \tilde{y}_t$.[9] Together with condition (79), we have, for all t,

$$\begin{aligned}
\tfrac{(\epsilon+1)\sigma\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{w}_t + \tfrac{\epsilon\sigma\Omega_2}{\epsilon\sigma+\Omega_1}\tilde{e}_t - \tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{t}_t &= \tfrac{\sigma\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{w}_t + \tfrac{\epsilon\sigma\Omega_2}{\epsilon\sigma+\Omega_1}\tilde{y}_t - \tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{g}_t \\
&= \tfrac{\epsilon\sigma(\Omega_1+\Omega_2)}{\epsilon\sigma+\Omega_1}\tilde{y}_t + \tfrac{\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{c}_t - \tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{g}_t \\
&= \tfrac{\epsilon\sigma(\Omega_1+\Omega_2)}{\epsilon\sigma+\Omega_1}\left[\tfrac{1}{\Omega_1+\Omega_2}\tilde{c}_t + \tfrac{\Omega_1+\Omega_2-1}{\Omega_1+\Omega_2}\tilde{g}_t\right] + \tfrac{\Omega_1}{\epsilon\sigma+\Omega_1}\tilde{c}_t - \tfrac{\epsilon\sigma(\Omega_1+\Omega_2-1)}{\epsilon\sigma+\Omega_1}\tilde{g}_t = \tilde{c}_t.
\end{aligned}$$

Substitute it into condition (78), we have

$$\tilde{c}_t = -\sigma\sum_{k=1}^{+\infty}\beta^{k-1}\bar{E}_t^c[\tilde{r}_{t+k}] + \tfrac{1-\beta}{\beta}\left\{\sum_{k=1}^{+\infty}\beta^k\bar{E}_t^c\left[\tilde{c}_{t+k}\right]\right\}. \tag{81}$$

This is exactly the same form of the consumption beauty contest, as condition (5).

Now let us state Proposition 8 formally here. Similar to Assumption 2, we assume $\tilde{g}_T = z + \eta$, where z and η are random variables, independent of one another and of any other shock in the economy, with $z \sim N(0, \sigma_z^2)$ and $\eta \sim N(0, \sigma_\eta^2)$. The former is realized at $t = 0$, and could be interpreted as news about government spending; the latter is realized at $t = T$ and is unpredictable prior to that point.

First consider the complete information outcome. Suppose z is commonly known starting at $t = 0$, we can find a scalar $\phi_{g,T}^*$ such that $\tilde{c}_0 - \tilde{c}_0^{trap} = \phi_{g,T}^* E_0[\tilde{g}_T]$, where $\tilde{c}_0^{trap}$ denotes the "liquidity trap" level of consumption (i.e., the one obtained when it is common knowledge that $\tilde{g}_T = 0$.) We have, when $\kappa > 0$,

$$\phi_{g,T}^* > 0, \text{ is strictly increasing in } T, \text{ and diverges to infinity as } T \to \infty. \tag{82}$$

Now consider the case in which z is not common knowledge. Similar to Section V, we consider the information structure specified in Assumption 3, in which let each agent receives a private signal about z at period 0. We can then find a scalar $\phi_{g,T}$ such that $\tilde{c}_0 - \tilde{c}_0^{trap} = \phi_{g,T}\bar{E}_0^c[\tilde{g}_T]$. We have, as long as $\kappa > 0$ and information is incomplete, that is $\lambda_c < 1$,[10]

$$\begin{aligned}
\phi_{g,T} \in \left(0, \phi_{g,T}^*\right), &\text{ is strictly increasing in } \lambda_c \text{ and } \lambda_f; \\
\text{the ratio } \phi_{g,T}/\phi_{g,T}^* &\text{ is strictly decreasing in } T \text{ and converges to 0 as } T \to \infty; \\
\text{finally, when } \lambda_c \text{ is sufficiently low, } &\phi_{g,T} \text{ also converges to 0 as } T \to \infty.
\end{aligned} \tag{83}$$

We start from the proof of condition (82). Similar to the proof Proposition 6, we can establish that there

[9]This expression is equivalent to $\frac{\Omega_2}{\Omega_1+\Omega_2}\tilde{e}_t + \frac{\Omega_1}{\Omega_1+\Omega_2}(\tilde{w}_t + \tilde{y}_t) = \frac{\Omega_2}{\Omega_1+\Omega_2}\tilde{e}_t + \frac{\Omega_1}{\Omega_1+\Omega_2}\left(\tilde{w}_t + \int_{\mathcal{I}_f}\tilde{l}_t^j dj\right) = \tilde{y}_t$. The last equation is true because $\frac{\Omega_1}{\Omega_1+\Omega_2}$ is steady state labor income to total income ratio (before deducting tax) and $\frac{\Omega_2}{\Omega_1+\Omega_2}$ is steady state dividend income to total income ratio (before deducting tax).

[10]For simplicity here, we always remove common knowledge about z among consumer here. We allow $\lambda_f \in (0, 1]$. In other words, we nest the case in which firms have perfect knowledge about z.

exists non-negative scalars $\left\{\phi^*_{g,\tau}, \varpi^*_{g,\tau}\right\}_{\tau\geq 0}$ such that, when z is commonly known, the equilibrium spending and inflation at any $t \leq T$ are given by

$$\tilde{c}_t - \tilde{c}_t^{trap} = \phi^*_{g,T-t} \cdot E_t[\tilde{g}_T], \tag{84}$$
$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} = \kappa\left(\Omega_c\left(\tilde{c}_t - \tilde{c}_t^{trap}\right) + (1-\Omega_c)\,\tilde{g}_t\right) + \varpi^*_{g,T-t} \cdot E_t[\tilde{g}_T], \tag{85}$$

where $\tilde{c}_t^{trap}$ and $\tilde{\pi}_t^{trap}$ denotes the "liquidity trap" level of consumption and inflation (i.e., the one obtained when it is common knowledge that $\tilde{g}_T = 0$.) Note that the Euler condition and NKPC with government spending under complete information can be written as:

$$\tilde{c}_t \quad = \quad -\sigma\left\{\tilde{R}_t - E_t\left[\tilde{\pi}_{t+1}\right]\right\} + E_t\left[\tilde{c}_{t+1}\right], \tag{86}$$
$$\tilde{\pi}_t \quad = \quad \kappa\left(\Omega_c\tilde{c}_t + (1-\Omega_c)\,\tilde{g}_t\right) + \beta E_t\left[\tilde{\pi}_{t+1}\right] + \varkappa\tilde{\mu}_t. \tag{87}$$

Using the above expressions, similar to the proof of Proposition 6, we can establish that $\phi^*_{g,0} = 0$, $\varpi^*_{g,0} = 0$, $\phi^*_{g,1} = \sigma\kappa\,(1-\Omega_c)$, $\varpi^*_{g,1} = \beta\kappa\,(1-\Omega_c)$, and for all $\tau \geq 1$,

$$\phi^*_{g,\tau+1} = (1+\sigma\kappa\Omega_c)\,\phi^*_{g,\tau} + \sigma\varpi^*_{g,\tau}, \tag{88}$$
$$\varpi^*_{g,\tau+1} = \beta\kappa\Omega_c\phi^*_{g,\tau} + \beta\varpi^*_{g,\tau}. \tag{89}$$

From condition (88), we can see when $\kappa > 0$, $\phi^*_{g,\tau}$ is positive, strictly increasing in τ and diverges to infinity. This proves condition (82). Similar to condition (59), one can also prove the following recursive relationship about ϕ^*_g:

$$\frac{\phi^*_{g,\tau+1}}{\phi^*_{g,\tau}} + \beta\frac{\phi^*_{g,\tau-1}}{\phi^*_{g,\tau}} = 1 + \beta + \sigma\kappa\Omega_c \quad \forall\tau \geq 1. \tag{90}$$

Moreover, as condition (61), we know $\frac{\phi^*_{g,\tau}}{\phi^*_{g,\tau-1}}$ must converge to $\Gamma^*_g > 0$, as $\tau \to \infty$:

$$\Gamma^*_g + \beta\frac{1}{\Gamma^*_g} = 1 + \beta + \sigma\kappa\Omega_c. \tag{91}$$

We now turn to the case of incomplete information and establish the proof of condition (83). Similar to the proof of Proposition 7, we can find $\phi_g, \varpi_g : (0,1] \times (0,1] \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that, for any $t \leq T-1$,

$$\tilde{c}_t - \tilde{c}_t^{trap} \quad = \quad \phi_g\left(\lambda_c, \lambda_f, T-t\right)\bar{E}^c_t[\tilde{g}_T], \tag{92}$$
$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} \quad = \quad \kappa\left(\Omega_c\left(\tilde{c}_t - \tilde{c}_t^{trap}\right) + (1-\Omega_c)\,\tilde{g}_t\right) + \varpi_g\left(\lambda_c, \lambda_f, T-t\right)\bar{E}^f_t[\tilde{g}_T]. \tag{93}$$

Using conditions (80) and (81), we have $\phi_g\left(\lambda_c, \lambda_f, 1\right) = \sigma\kappa\,(1-\Omega_c)$, $\varpi_g\left(\lambda_c, \lambda_f, 1\right) = \beta\kappa\,(1-\Omega_c)$ and, for all $\tau \geq 2$,

$$\phi_g(\lambda_c, \lambda_f, \tau) = \sigma\beta^{\tau-1}(1-\Omega_c)\kappa + \sum_{k=1}^{\tau-1} \beta^{k-1}\left[(1-\beta+\sigma\kappa\Omega_c)\lambda_c\phi_g(\lambda_c, \lambda_f, \tau-k) + \sigma\lambda_f\varpi_g(\lambda_c, \lambda_f, \tau-k)\right]; \tag{94}$$

$$\varpi_g(\lambda_c, \lambda_f, \tau) = (1-\Omega_c)\frac{\kappa}{\theta}(\beta\theta)^{T-t} + \sum_{k=1}^{\tau-1}(\beta\theta)^k\left(\tfrac{\kappa\lambda_c\Omega_c}{\theta}\phi_g(\lambda_c, \lambda_f, \tau-k) + \tfrac{(1-\theta)\lambda_f}{\theta}\varpi_g(\lambda_c, \lambda_f, \tau-k)\right). \tag{95}$$

Together, we can establish, for all $\tau \geq 2$,

$$\phi_g(\lambda_c, \lambda_f, \tau) = (\beta + (1-\beta+\sigma\kappa\Omega_c)\lambda_c)\phi_g(\lambda_c, \lambda_f, \tau-1) + \sigma\lambda_f\varpi_g(\lambda_c, \lambda_f, \tau-1); \tag{96}$$

$$\varpi_g(\lambda_c, \lambda_f, \tau) = \kappa\beta\lambda_c\Omega_c\phi_g(\lambda_c, \lambda_f, \tau-1) + \beta\left[\theta + (1-\theta)\lambda_f\right]\varpi_g(\lambda_c, \lambda_f, \tau-1). \tag{97}$$

From the above conditions, we can see that for for all $\tau \geq 2$, $\phi_{g,\tau} = \phi_g(\lambda_c, \lambda_f, \tau)$ is strictly increasing in λ_c and λ_f. As $\phi^*_{g,\tau} = \phi_g(1, 1, T)$, we also have $\phi_{g,\tau} \in \left(0, \phi^*_{g,\tau}\right)$.

Now, we now prove that, whenever $\lambda_c < 1$, the ratio $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}} = \frac{\phi_g(\lambda_c, \lambda_f, \tau)}{\phi^*_{g,\tau}}$ is strictly decreasing in $\tau \geq 1$ and converges to 0 as $\tau \to \infty$. To this goal, similar to condition (74), we try to establish that

$$\frac{\phi_{g,\tau}}{\phi_{g,\tau-1}} + \beta\frac{\phi_{g,\tau-2}}{\phi_{g,\tau-1}} \leq 1 + \beta + \sigma\kappa\Omega_c\lambda_c < 1 + \beta + \sigma\kappa\Omega_c \quad \forall\tau \geq 3. \tag{98}$$

From condition (96), we have, for $\tau \geq 3$,

$$\phi_{g,\tau} = (\beta + (1-\beta)\lambda_c)\phi_{g,\tau-1} + \sigma\kappa\Omega_c\lambda_c\phi_{g,\tau-1} + \sigma\lambda_f\varpi_{g,\tau-1},$$
$$\tfrac{\beta}{\beta+(1-\beta)\lambda_c}\phi_{g,\tau-1} = \beta\phi_{g,\tau-2} + \tfrac{\sigma\beta\kappa\Omega_c\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{g,\tau-2} + \tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{g,\tau-2}.$$

From the previous two conditions, we have, for $\tau \geq 3$,

$$\begin{aligned}\phi_{g,\tau} + \beta\phi_{g,\tau-2} = {} & (\beta + (1-\beta)\lambda_c)\phi_{g,\tau-1} + \sigma\kappa\Omega_c\lambda_c\phi_{g,\tau-1} + \sigma\lambda_f\varpi_{g,\tau-1} \\ & + \tfrac{\beta}{\beta+(1-\beta)\lambda_c}\phi_{g,\tau-1} - \tfrac{\sigma\beta\kappa\Omega_c\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{g,\tau-2} - \tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{g,\tau-2}.\end{aligned} \tag{99}$$

Note that, for $\tau \geq 3$ and $\lambda_c, \lambda_f \in (0, 1]$, we have

$$\left[(\beta + (1-\beta)\lambda_c) + \sigma\kappa\Omega_c\lambda_c + \tfrac{\beta}{\beta+(1-\beta)\lambda_c}\right]\phi_{g,\tau-1} \leq (1 + \beta + \sigma\kappa\Omega_c\lambda_c)\phi_{g,\tau-1},$$

and from condition (97), we have

$$
\begin{aligned}
&\sigma\lambda_f\varpi_{g,\tau-1}-\tfrac{\sigma\beta\kappa\Omega_c\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{g,\tau-2}-\tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{g,\tau-2}\\
=&\sigma\lambda_f\left(\kappa\beta\lambda_c\Omega_c\phi_{g,\tau-2}+\beta\left[\theta+(1-\theta)\lambda_f\right]\varpi_{g,\tau-2}\right)-\tfrac{\sigma\beta\kappa\Omega_c\lambda_c}{\beta+(1-\beta)\lambda_c}\phi_{g,\tau-2}-\tfrac{\sigma\beta\lambda_f}{\beta+(1-\beta)\lambda_c}\varpi_{g,\tau-2}\\
=&\sigma\kappa\beta\Omega_c\lambda_c\left(\lambda_f-\tfrac{1}{\beta+(1-\beta)\lambda_c}\right)\phi_{g,\tau-2}+\sigma\beta\lambda_f\left[\theta+(1-\theta)\lambda_f-\tfrac{1}{\beta+(1-\beta)\lambda_c}\right]\varpi_{g,\tau-2}\\
\leq&0.
\end{aligned}
$$

Together with condition (99), we reach at condition (98). To prove $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}$ is strictly decreasing in τ, note that we already prove that $\frac{\phi_{g,2}}{\phi^*_{g,2}}<1=\frac{\phi_{g,1}}{\phi^*_{g,1}}$. We proceed by induction on τ. If $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}<\frac{\phi_{g,\tau-1}}{\phi^*_{g,\tau-1}}$ for $\tau\geq 2$, we have $\frac{\phi_{g,\tau-1}}{\phi_{g,\tau}}>\frac{\phi^*_{g,\tau-1}}{\phi^*_{g,\tau}}$. From (90) and (98), we have $\frac{\phi_{g,\tau+1}}{\phi_{g,\tau}}<\frac{\phi^*_{g,\tau+1}}{\phi^*_{g,\tau}}$ and thus $\frac{\phi_{g,\tau+1}}{\phi^*_{g,\tau+1}}<\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}$. This finishes the proof that $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}$ is strictly decreasing in $\tau\geq 1$.

To prove that $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}$ converges to 0 as $\tau\to\infty$. Because $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}>0$ is strictly decreasing in $\tau\geq 1$, there exists $\Gamma_g\in[0,1)$ such that $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}\to\Gamma_g$ as $\tau\to\infty$. If $\Gamma_g>0$, we have $\frac{\phi_{g,\tau}}{\phi^*_{g,\tau}}\frac{\phi^*_{g,\tau-1}}{\phi_{g,\tau-1}}\to 1$ as $\tau\to\infty$. Because $\frac{\phi^*_{g,\tau}}{\phi^*_{g,\tau-1}}\to\Gamma^*_g$, we have $\frac{\phi_{g,\tau}}{\phi_{g,\tau-1}}\to\Gamma^*_g$ and $\frac{\phi_{g,\tau-2}}{\phi_{g,\tau-1}}\to\frac{1}{\Gamma^*_g}$ as $\tau\to\infty$. From condition (91), we have $\frac{\phi_{g,\tau}}{\phi_{g,\tau-1}}+\beta\frac{\phi_{g,\tau-2}}{\phi_{g,\tau-1}}\to 1+\beta+\sigma\kappa\Omega_c$ as $\tau\to\infty$. However, this is inconsistent with (98) as $\lambda_c<1$ and $\kappa>0$. As a result, $\Gamma_g=0$.

Finally, we prove that, when λ_c is sufficiently low, $\phi_g(\lambda_c,\lambda_f,\tau)$ converges to zero as $\tau\to\infty$. The eigenvalues of the dynamic system $(\phi_{g,\tau},\varpi_{g,\tau})$ based on conditions (96) and (97) are

$$
\begin{aligned}
m_1=&\frac{\beta+(1-\beta+\sigma\kappa\Omega_c)\lambda_c+\beta\left[(1-\theta)\lambda_f+\theta\right]}{2}\\
&-\frac{\sqrt{\left(\beta+(1-\beta+\sigma\kappa\Omega_c)\lambda_c-\beta\left[(1-\theta)\lambda_f+\theta\right]\right)^2+4\sigma\beta\lambda_f\lambda_c\Omega_c\kappa}}{2}>0,\\
m_2=&\frac{\beta+(1-\beta+\sigma\kappa\Omega_c)\lambda_c+\beta\left[(1-\theta)\lambda_f+\theta\right]}{2}\\
&+\frac{\sqrt{\left(\beta+(1-\beta+\sigma\kappa\Omega_c)\lambda_c-\beta\left[(1-\theta)\lambda_f+\theta\right]\right)^2+4\sigma\beta\lambda_f\lambda_c\Omega_c\kappa}}{2}>m_1.
\end{aligned}
$$

Note that $\lim_{\lambda_c\to 0}m_2=\beta<1$. As a result, when λ_c is sufficiently low, both eigenvalues are below 1, which means that $\phi_g(\lambda_c,\lambda_f,\tau)$ converges to zero as $\tau\to\infty$.

Appendix B. Rational Inattention and Learning

In this appendix, we first sketch how the friction we consider can be recast as the product of rational inattention. We next extend Theorem 1 to two leading forms of learning studied in the literature. We finally prove an asymptotic version of our horizon effect for arbitrary forms of learning.

The Friction as the Product of Rational Inattention. We now briefly sketch how the friction we consider can be recast as the product of rational inattention and, in this sense, a form of bounded rationality.[11]

We let Θ_T be Normally distributed and, to sharpen the exposition, we assume that Θ_T is realized at $t = 0$ (think of Θ_T as being itself the news). The typical agent is nevertheless unable to observe Θ_T perfectly. Instead, for every t, the action $a_{i,t}$ must be measurable in $\omega_i^t \equiv \{\omega_{i,\tau}\}_{\tau \leq t}$, where $\omega_{i,\tau}$ is a noisy signal obtained in period τ. The noise is assumed to be independent across the agent.[12] This guarantees that all aggregate outcomes are functions of Θ_T and, therefore, we can reduce the rational-inattention problem faced by each agent to the choice of a sequence of signals about Θ_T. We finally let these signals be chosen optimally, that is, so as to maximize the agent's ex ante payoff, subject to the following constraint:

$$\mathcal{I}\left(\, \omega_{i,t}, \Theta_T \mid \omega_i^{t-1} \,\right) \leq \kappa^{RI}, \tag{100}$$

where $\mathcal{I}\left(\omega_{i,t}, \Theta_T | \omega_i^{t-1}\right)$ is the (entropy-based) information flow between the period-t signal and Θ_T, conditional on the agent's past information, and $\kappa^{RI} > 0$ is an exogenous scalar.

The usual interpretation of constraint (100) is that it captures the agent's limited cognitive capacity in tracking Θ_T. But since beliefs about Θ_T map, in equilibrium, to beliefs of future outcomes, one can also think of (100) as a constraint on the agent's ability to figure out the likely effects of the underlying variation in Θ_T. This echoes Tirole (2015), who interprets rational inattention in games as a form of "costly contemplation."

As long as the prior about Θ_T is Gaussian and the objective function is quadratic, which is the case here by assumption, the optimal signal is also Gaussian. Furthermore, the noise in the signal has to be independent across periods, or else the agent could economize on cognitive costs, that is, relax the constraint in (100). These arguments are standard; see, e.g., Mackowiak, Matejka and Wiederholt (2017). The case studied here is actually far simpler than the one studied in the literature, because the relevant fundamental (Θ_T) does not vary as time passes. We infer that the optimal signal at every $t \leq T-1$ is given by $\omega_{i,t} = \Theta_T + v_{i,t}$, where the noise $v_{i,t}$ is orthogonal to both Θ_T and $\{v_{i,\tau}\}_{\tau<t}$. Letting τ_t denotes the precision (i.e., the reciprocal of the variance) of this noise, we have that the period-t information flow is given by

$$\mathcal{I}\left(\, \omega_{i,t}, \Theta_T \mid \omega_i^{t-1} \,\right) = \tfrac{1}{2}\log_2\left(1 + \frac{\tau_t}{\varsigma_t}\right),$$

[11] For the rational inattention problem to be well-defined, we need to specify a payoff structure behind condition (11). For example, we can think player i's payoff is $\mathcal{U}_i = \sum_t \beta^t U(a_{i,t}, a_{i,t+1}; \Theta_t, a_t)$, where U is a reverse-engineered quadratic utility function so that the player's best-response condition is given by (11).

[12] Although this assumption is separate the information-flow constraint (100), it is standard in the literature (e.g., Woodford, 2003, Mackowiak and Wiederholt, 2009, Luo et al., 2017) and seems appealing if one interprets the noise as the product of cognitive limitations. It is also broadly consistent with experimental evidence (e.g., Khaw, Stevens and Woodford, 2016).

where ς_t denotes the precision of the agent's prior in the beginning of period t; the latter is defined recursively by $\varsigma_0 = \sigma_\theta^{-2}$ and $\varsigma_{t+1} = \varsigma_t + \tau_t$ for $t \in \{0, ..., T-1\}$. It follows that the information constraint (100) pins down the sequence $\{\tau_t\}_{t=0}^{T-1}$ as a function of κ^{RI} and σ_θ^2 alone. All in all, the setting we have considered here is therefore nested in the cases with learning studied in the next part of this appendix, for which Theorem 1 applies. This completes the rational-inattention interpretation of our results.[13]

The Horizon Effect with Sticky Information or Noisy Private Learning. We now extend Theorem 1 to two leading forms of learning studied in the literature. Let $\Theta_T \sim N\left(0, \sigma_\theta^2\right)$ and consider the following two cases of learning.

Case 1. Agents become gradually aware of Θ_T, as in Mankiw and Reis (2002) and Wiederholt (2015). In particular, at each $t \in \{0, ..., T-1\}$, a fraction λ_{sticky} of agents who have not become aware about Θ_T become aware about Θ_T. Θ_T becomes commonly known at period T.

Case 2. Agents receive a new private signal each period, as in Woodford (2003), Nimark (2008), and Mackowiak and Wiederholt (2009). In particular, at at each $t \in \{0, ..., T-1\}$, agent i's new information about Θ_T is summarized in the private signal $s_{i,t} = \Theta_T + v_{i,t}$, where $v_{i,t} \sim N\left(0, \sigma_{v,t}^2\right)$ is i.i.d across i and t, and independent of Θ_T. Θ_T becomes commonly known at period T.

In both cases, there exists $\{\lambda_t\}_{t=0}^{T-1}$ such that, for all t, $\lambda_t \in (0,1)$ and, for any $h \in \{1, ..., T\}$ and $0 \le t_1 < t_2 < \cdots < t_h < T$,

$$\bar{E}_{t_1}[\bar{E}_{t_2}[...[\bar{E}_{t_h}[\Theta_T]...]] = \lambda_{t_1} \cdots \lambda_{t_h} \Theta_T, \tag{101}$$

In case 1, $\lambda_t = 1 - (1 - \lambda_{sticky})^{t+1}$. In case 2, $\lambda_t = \frac{\sum_{\tau=0}^{t} \sigma_{v,\tau}^{-2}}{\sum_{\tau=0}^{t} \sigma_{v,\tau}^{-2} + \sigma_\theta^{-2}}$.

Now we prove by induction that, for all $t \le T-1$,

$$a_t = (\gamma + \alpha) \left\{ \Pi_{\tau=t+1}^{T-1} (\gamma + \lambda_\tau \alpha) \bar{E}_t[\Theta_T] \right\}. \tag{102}$$

Since $\Theta_t = 0$ for all $t \neq T$, together with condition (16), we have $a_T = \Theta_T$ and $a_{T-1} = (\gamma + \alpha) \bar{E}_{T-1}[\Theta_T]$. As a result, condition (102) holds for $t = T-1$. Now, pick a $t \le T-2$, assume that the claim holds for all $\tau \in \{t+1, ..., T-1\}$, and let us prove that it also holds for t. Using the claim for all $\tau \in \{t+1, ..., T-1\}$,

[13] Two remarks are worth making. First, suppose that the agents have a limited cognitive capacity to allocate, not per period, but across the entire horizon. In this case, the series of per-period constraints seen in condition (100) are replaced by a single constraint over the entire horizon, namely $\mathcal{I}\left(\{\omega_{i,t}\}_{t=0}^{T-1}, \Theta_T\right) \le \kappa^{RI}$. It then becomes optimal to allocate all capacity to the period-0 signal, which means that this case can justify to our baseline analysis, which assumes away learning. Second, suppose that the news about the fundamental of interest (say, monetary policy) come at the same time with news about another fundamental (say, TFP) and that the agents can economize on cognitive effort by obtaining a joint signal of all the news. In this case, rational inattention can contribute to confounding of one kind of news with another, a scenario not considered here.

condition (16), and condition (101), we have, for $t \leq T-2$,

$$a_t = \gamma^{T-t}\bar{E}_t\left[\Theta_T\right] + \alpha\bar{E}_t\left[a_{t+1}\right] + \alpha\sum_{k=2}^{T-t}\gamma^{k-1}\bar{E}_t\left[a_{t+k}\right];$$

$$\gamma\bar{E}_t\left[a_{t+1}\right] = \lambda_{t+1}\gamma^{T-t}\bar{E}_t\left[\Theta_T\right] + \alpha\lambda_{t+1}\sum_{k=2}^{T-t}\gamma^{k-1}\bar{E}_t\left[a_{t+k}\right].$$

As a result, we have $a_t = \left(\frac{\gamma}{\lambda_{t+1}} + \alpha\right)\bar{E}_t\left[a_{t+1}\right]$. Together with condition (102) for $t+1$, we have

$$a_t = (\gamma+\alpha)\left\{\Pi_{\tau=t+1}^{T-1}\left(\gamma+\lambda_\tau\alpha\right)\bar{E}_t[\Theta_T]\right\}.$$

This proves condition (102) for all $t \leq T-1$. As a result, $\phi_T = (\gamma+\alpha)\,\Pi_{t=1}^{T-1}\left(\gamma+\lambda_t\alpha\right)$. Together with the fact that $\lambda_t \in (0,1)$ and $\phi_T^* = (\gamma+\alpha)^T$, we prove Theorem 1 for the case with learning.

The Limit Property with Arbitrary Learning. As noted in the main text, it is possible to prove, under quite general conditions, an *asymptotic* version of our horizon effect: as long as the higher-order uncertainty is bounded away from zero (in a sense we make precise now), the scalar ϕ_T becomes vanishingly small relative to ϕ_T^* as $T \to \infty$.

For any $t \leq T-1$ and any $k \in \{1, ..., T-t\}$, we henceforth let B_t^k collect all the relevant k-order beliefs, as of period t :

$$B_t^k \equiv \left\{x \;:\; \exists(t_1, t_2, \cdots, t_k),\ \text{with } t = t_1 < t_2 < \cdots < t_k \leq T-1,\ \text{such that } x = \bar{E}_{t_1}[\bar{E}_{t_2}[\cdots\bar{E}_{t_k}[\Theta_T]\cdots]]\right\}.$$

We next introduce the following assumption.

Assumption 4 (Non-Vanishing Higher-Order Uncertainty) *There exists an $\epsilon > 0$ such that:*

(i) For all $t \in \{0, ..., T-1\}$, there exists at least a mass ϵ of agents such that

$$Var\left(E_t[x]|\omega_i^t\right) \geq \epsilon Var\left(E_t\left[x\right]\right),$$

for all $x \in B_\tau^k \cup \{\Theta_T\}$, $\tau \in \{t+1, ..., T-1\}$, and $k \in \{1, ..., T-\tau\}$, where ω_i^t summarizes agent i's information at period t and $E_t\left[x\right]$ denotes the rational expectation of variable x conditional on the union of information of all agents in the economy available at period t.

(ii) $Var\left(\bar{E}_0[\Theta_T]\right) \geq \epsilon$.

To interpret this assumption, note that complete information imposes that $E_t\left[x\right]$ is known to every agent, and therefore that $Var\left(E_t\left[x\right]]|\omega_i^t\right) = 0$, regardless of how volatile $E_t\left[x\right]$ itself is. By contrast, letting $Var\left(E_t\left[x\right]|\omega_i^t\right) > 0$ whenever $Var\left(E_t\left[x\right]\right) > 0$ is essentially *tautological* to assuming that agents have incomplete information or, equivalently, that they face higher-order uncertainty. Relative to this tautology, part (i) introduces an arbitrarily small bound on the level of higher-order uncertainty. This bound guarantees that the higher-order uncertainty does not vanish as we let T go to infinity. Part (ii), on the other hand, means simply that there is non-trivial variation in first-order beliefs in the first place. The next result then formalizes our point that our horizon effect, at least in its limit form, holds for arbitrary forms of learning.

Proposition 9 (Limit) *Under Assumption 4, the ratio $\frac{\phi_T}{\phi_T^*}$ converges to zero as $T \to \infty$.*

Proof of Proposition 9. We first prove that, under Assumption (4),

$$Var\left(y\right) \leq \left(1-\epsilon^2\right)^k Var\left(\Theta_T\right). \tag{103}$$

for any $t \leq T-1$ and $y = \bar{E}_t \bar{E}_{t_2} ... \bar{E}_{t_k}[\Theta_T] \in B_t^k$.

To simplify notation, let $x = \bar{E}_{t_2} ... \bar{E}_{t_k}[\Theta_T]$ for $k \geq 2$, and $x = \Theta_T$ for $k = 1$. From Assumption 4, we have, there is at least a mass ϵ of agents such that

$$Var\left(E_t\left[x\right]|\omega_i^t\right) \geq \epsilon Var\left(E_t\left[x\right]\right).$$

As a result,

$$E\left[Var\left(E_t\left[x\right]|\omega_i^t\right)|\Omega_t\right] \geq \epsilon^2 Var\left(E_t\left[x\right]\right),$$

where Ω_t is the cross-sectional distribution of information ω_i^t at period t. Using the law of total variance, we have

$$Var\left(E_t\left[x\right]\right) = E\left[Var\left(E_t\left[x\right]|\omega_i^t\right)|\Omega_t\right] + Var\left(E\left[E_t\left[x\right]|\omega_i^t\right]\right) = E\left[Var\left(E_t\left[x\right]|\omega_i^t\right)|\Omega_t\right] + Var\left(E\left[x|\omega_i^t\right]\right).$$

As a result, we have[14]

$$\begin{aligned} Var\left(y\right) = Var\left(\bar{E}_t \bar{E}_{t_2} ... \bar{E}_{t_k}[\Theta_T]\right) = Var\left(\bar{E}_t\left[x\right]\right) \leq Var\left(E\left[x|\omega_i^t\right]\right) \leq \left(1-\epsilon^2\right) Var\left(E_t\left[x\right]\right) \\ \leq \left(1-\epsilon^2\right) Var\left(x\right) = \left(1-\epsilon^2\right) Var\left(\bar{E}_{t_2} ... \bar{E}_{t_k}[\Theta_T]\right). \end{aligned}$$

Iterating the previous condition proves (103).

Condition (103) provides an upper bound for the variance of all k-th order belief. Together with the fact that $\phi_T^* = s_{T,T}$ and, for any random variables X, Y and scalars $a, b \geq 0$,

$$\begin{aligned} Var(aX+bY) &= a^2 Var(X) + 2ab Cov\left(X,Y\right) + b^2 Var\left(Y\right) \\ &\leq a^2 Var(X) + 2ab\sqrt{Var\left(X\right) Var\left(Y\right)} + b^2 Var\left(Y\right) \\ &= \left(a\sqrt{Var\left(X\right)} + b\sqrt{Var\left(Y\right)}\right)^2, \end{aligned} \tag{104}$$

[14] We use the fact that for any random variable X, and any information I, $Var\left(E\left[X|I\right]\right) \leq Var\left(X\right)$. We also use the fact that $\bar{E}_t\left[\cdot\right] = E\left[E\left[\cdot|\omega_i^t\right]|\Omega_t\right]$, where Ω_t is the cross sectional distribution of ω_i^t at time t,

we have

$$\begin{aligned}
\left(\frac{\phi_T}{\phi_T^*}\right)^2 &= \left(\frac{Cov\left(a_0, \bar{E}_0\left[\Theta_T\right]\right)}{\phi_T^* Var\left(\bar{E}_0\left[\Theta_T\right]\right)}\right)^2 \leq \frac{Var\left(a_0\right)}{\left[\phi_T^*\right]^2 Var\left(\bar{E}_0[\Theta_T]\right)} \\
&\leq \frac{1}{Var\left(\bar{E}_0[\Theta_T]\right)} \left[\sum_{k=1}^{T}\left(\frac{\chi_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}}\sqrt{Var\left(\Theta_T\right)}\right)\right]^2 \\
&= \left[\sum_{k=1}^{T}\left(\frac{\chi_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}}\right)\right]^2 \frac{Var\left(\Theta_T\right)}{Var\left(\bar{E}_0[\Theta_T]\right)}. \qquad (105)
\end{aligned}$$

Further note that, for any $\vartheta > 0$, there exists $h \in \mathbb{N}_+$ such that $\frac{\left(1-\epsilon^2\right)^{\frac{h}{2}}}{1-\left(1-\epsilon^2\right)^{\frac{1}{2}}} \leq \frac{\vartheta}{2}$. From Theorem 2, there exists $T^* \in \mathbb{N}_+$ such that, for all $T \geq T^*$, $\sum_{k=1}^{h-1}\frac{\chi_{k,T}}{s_{T,T}} \leq \frac{\vartheta}{2}$. As a result, for all $T \geq \max\{T^*, h\}$,

$$\sum_{k=1}^{T}\frac{\chi_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}} \leq \sum_{k=1}^{h-1}\frac{\chi_{k,T}}{s_{T,T}} + \sum_{k=h}^{T}\left(1-\epsilon^2\right)^{\frac{k}{2}} \leq \frac{\vartheta}{2} + \frac{\left(1-\epsilon^2\right)^{\frac{h}{2}}}{1-\left(1-\epsilon^2\right)^{\frac{1}{2}}} \leq \vartheta.$$

This proves

$$\sum_{k=1}^{T}\left(\frac{\chi_{k,T}}{s_{T,T}}\left(1-\epsilon^2\right)^{\frac{k}{2}}\right) \to 0 \text{ as } T \to +\infty.$$

Together with (105) and the fact that $Var\left(\bar{E}_0[\Theta_T]\right) \geq \epsilon$, the proof of Proposition 9 is completed.

Appendix C. Additional Results for the New Keynesian Model

In this Appendix, we provide a few additional results regarding the application of our insights in the context of a liquidity trap. We first explain how our results regarding the forward-guidance puzzle can be understood under the lenses of a discounted Euler condition and a discounted NKPC, and draw certain connections to the literature. We next show how our insights help lessen the paradox of flexibility. We finally show that all the results of Section IV extend to the new type of beauty contest seen in condition (22).

Discounted Euler Condition and Discounted NKPC. Proposition 5 has already indicated how the lack of common knowledge is akin to introducing additional discounting in the forward-looking equations of a macroeconomic model. We now illustrate how this helps recast our results regarding forward guidance and fiscal multipliers under the lenses of a discounted Euler condition and a discounted NKPC.

For the present purposes, we make a minor modification to the setting used in Section V: for $t \leq T$, we let the firms lack knowledge of the concurrent level of marginal cost. For simplicity, we also let the firms and the consumers face the same level of friction, that is, we set $\lambda_c = \lambda_f = \lambda$. These modifications are not strictly needed but sharpen the representation offered below.[15]

Proposition 10 *The power of forward guidance in the absence of common knowledge, ϕ_T, is the same as that in a representative-agent variant in which the Euler condition and the NKPC are modified as follows, for all $t \leq T-1$:*

$$\tilde{y}_t = -\sigma \left\{ \tilde{R}_t - \lambda E_t \left[\tilde{\pi}_{t+1} \right] \right\} + M_c E_t \left[\tilde{y}_{t+1} \right] \quad (106)$$

$$\tilde{\pi}_t = \kappa' \tilde{y}_t + \beta M_f E_t \left[\tilde{\pi}_{t+1} \right] + \varkappa \tilde{\mu}_t, \quad (107)$$

where $M_c \equiv \beta + (1-\beta)\lambda \in (\beta, 1]$, $M_f \equiv \theta + (1-\theta)\lambda \in (\theta, 1]$, and $\kappa' \equiv \kappa\lambda$.[16]

This result, which is analogous to Proposition 5 in our abstract setting, maps the incomplete-information ϕ_T of the economy under consideration to the complete-information ϕ_T^* of a *variant* economy, in which the Euler condition and the NKPC have been "discounted" in the manner described above. When we remove common knowledge, it is *as if* the representative consumer discounts her expectations of next period's aggregate income and inflation by a factor equal to, respectively, M_c and λ; and it is *as if* the representative firm discounts the future inflation by a factor equal to M_f.[17]

Consider first the Euler condition. When β is close to 1, the discount on future consumption, M_c, is close to 1, even if λ is close to zero. This underscores that the multiplier inside the demand block—which gets attenuated by the absence of common knowledge—is weak in the textbook version of the New Key-

[15]Without these modifications, the obtained representation is a bit less elegant, but the essence remains the same; see Proposition 10 in the first NBER version of our paper, Angeletos and Lian (2016).

[16]To be precise, condition (106) holds with $M_c = 1$ for $t = T-1$.

[17]The change in the slope of the NKPC, from κ to κ', is of relative little interest to us, because the effect of the informational friction through this slope cannot be identified separately from that of a higher Frisch elasticity or less steep marginal costs.

nesian model. As mentioned in the main text, short horizons, counter-cyclical precautionary savings, and feedback effects between housing prices and consumer spending tend to reinforce this multiplier, thereby also increasing the discounting caused by the absence of common knowledge. Also note that, while future consumption is discounted by M_c, future inflation is discounted by λ. Clearly, this can have a significant effect on the joint dynamics of spending and inflation even when M_c is close to 1.

Consider next the NKPC. For the textbook parameterization of the degree of price stickiness (meaning a price revision rate, $1-\theta$, equal to $1/3$), the effective discount factor, M_f, falls from 1 to $.9$ as we move from common knowledge ($\lambda = 1$) to the level of imperfection assumed in our numerical example ($\lambda = .75$). The magnitude of this discount helps explain the sizable effects seen in Figure 1. Under the parameterization we consider, the *actual* response of inflation to news about future demand is greatly reduced relative the common-knowledge benchmark. The fact that the average consumer underestimates the inflation response, as well as the spending of other consumers, reinforces this effect and helps further attenuate the feedback loop between inflation and spending.

Relation to Gabaix (2016) and McKay, Nakamura and Steinsson (2017). Related forms of discounting appear in McKay, Nakamura and Steinsson (2017), for the Euler condition, and in Gabaix (2016), for both the Euler condition and the NKPC. In this regards, these papers and ours are complementary to one another. However, the underlying theory and its empirical manifestations are different.

McKay, Nakamura and Steinsson (2017) obtain a discounted Euler condition at the aggregate level by introducing a specific combination of heterogeneity and market incompleteness that forces some agents to hit their borrowing constraints and breaks the individual-level Euler condition. This theory therefore ties the resolution of forward guidance to microeconomic evidence about the response of individual consumption to idiosyncratic shocks. By contrast, our theory ties the resolution of forward guidance to survey evidence about the response of average forecast errors to the underlying policy news. The two theories can therefore be quantified independently from one another—and it's an open question which is one is more relevant in the context of forward guidance.

Gabaix (2016) on the other hand, assumes two kinds of friction. The first is that agents are less responsive to any variation in interest rates and incomes due to "sparsity" (a form of adjustment cost). The second is that agents underestimate the response of future *aggregate* outcomes to exogenous shocks. The first is of purely decision-theoretic nature and, as the one in McKay, Nakamura and Steinsson (2017), amounts to a distortion of the individual-level Euler condition. The second is more closely related to the one we have obtained here: by anchoring expectations of *aggregate* outcomes, it gives rise to discounting only at the aggregate level. In this regard, Gabaix's theory and ours have a similar empirical implication: they both let the average forecast of future inflation and income respond less than the complete-information, rational-expectations, benchmark. Yet, our theory makes the following distinct prediction, which is consistent with the evidence in Coibion and Gorodnichenko (2012): the forecast errors, and the associated discounting, ought to decrease as time passes, agents accumulate more information, and higher-order beliefs converge to first-order beliefs.

Proof of Proposition 10. Let us first focus on the incomplete-information ϕ_T. When firms lack common knowledge of the concurrent level of marginal cost, condition (62) continues to hold but condition (63) becomes, for any $t \leq T$,

$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} = \kappa \sum_{k=0}^{T-t} (\beta\theta)^k \bar{E}_t^f \left[\tilde{y}_{t+k} - \tilde{y}_{t+k}^{trap}\right] + \tfrac{1-\theta}{\theta} \sum_{k=1}^{T-t} (\beta\theta)^k \bar{E}_t^f \left[\tilde{\pi}_{t+k} - \tilde{\pi}_{t+k}^{trap}\right].$$

Slightly different from conditions (64) and (65),[18] we can find functions $\phi, \omega : (0,1] \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that, for any $t \leq T-1$,

$$\tilde{y}_t - \tilde{y}_t^{trap} = -\phi(\lambda, T-t) \bar{E}_t^c[\tilde{R}_T], \tag{108}$$

$$\tilde{\pi}_t - \tilde{\pi}_t^{trap} = -\omega(\lambda, T-t) \bar{E}_t^f[\tilde{R}_T], \tag{109}$$

where $\phi(\lambda, 1) = \sigma(1+\sigma\lambda\kappa)$, $\omega(\lambda, 1) = \kappa\lambda\sigma(1+\sigma\lambda\kappa) + \kappa\beta(\theta + (1-\theta)\lambda)\sigma$, and for $t \leq T-2$,

$$\phi(\lambda, T-t) = (\beta + (1-\beta)\lambda) \phi(\lambda, T-t-1) + \sigma\lambda\omega(\lambda, T-t-1), \tag{110}$$

$$\omega(\lambda, T-t) = \beta(\theta + (1-\theta)\lambda) \omega(\lambda, T-t-1) + \kappa\lambda\phi(\lambda, T-t). \tag{111}$$

We now derive the complete-information ϕ_T^* and ω_T^* of a *variant* economy, where they denote how the output and inflation at $t=0$ responds to shocks to the representative agent's belief about $\tilde{R}_T$ at $t=0$. From conditions (106), (107) and footnote 16 in the appendix, we have $\phi_1^* = \sigma(1+\sigma\lambda\kappa)$, $\omega_1^* = \kappa\lambda\sigma(1+\sigma\lambda\kappa) + \kappa\beta(\theta + (1-\theta)\lambda)\sigma$, and, for $t \leq T-2$,

$$\phi_{T-t}^* = (\beta + (1-\beta)\lambda) \phi_{T-t-1}^* + \sigma\lambda\omega_{T-t-1}^*, \tag{112}$$

$$\omega_{T-t}^* = \beta(\theta + (1-\theta)\lambda) \omega_{T-t-1}^* + \kappa\lambda\phi_{T-t}^*. \tag{113}$$

The previous conditions coincide with conditions (110) and (111), and prove Proposition 10.

On the Paradox of Flexibility. We now consider the implications of our insights for the paradox of flexibility. In the standard model, the power of forward guidance and the fiscal multiplier vis-a-vis future government spending increase with the degree of price flexibility: ϕ_T increases with κ.[19] This property is directly related to the "paradox of flexibility" (Eggertsson and Krugman, 2012). The next result proves, in

[18] There are two differences compared to conditions (64) and (65). First, as we impose $\lambda_c = \lambda_f = \lambda$ here, ϕ and ω are functions of λ, the common parameter characterizing the degree of information friction. Second, as firms lack common knowledge of the concurrent level of marginal cost, it is easier to let ω measure how inflation as a whole responds to the average firm's belief about $\tilde{R}_T$.

[19] Whenever we vary κ, we vary θ while keeping the Frisch elasticity constant, which means that variation in κ maps one-to-one to variation in the degree of price flexibility.

code/figure2.eps

Figure 2: Varying the degree of price flexibility.

effect, that the mechanism identified in our paper helps diminish this paradox as well.

Proposition 11 (Price Flexibility) *Let ϕ_T be the scalar characterized in either Proposition 7 or Proposition 8 and set $\lambda_f = 1$. We have $\frac{\partial \phi_T}{\partial \kappa} > 0$ and $\frac{\partial}{\partial \lambda_c}\left(\frac{\partial \phi_T}{\partial \kappa}\right) > 0$. That is, the power of forward guidance and the fiscal multiplier vis-a-vis future government spending increase with the degree of price flexibility, but at a rate that is slower the greater the departure from common knowledge.*

This finding is an example of how lack of common knowledge reduces the paradox of flexibility more generally. In the standard model, a higher degree of price flexibility raises the GE effects of all kinds of demand shocks—whether these come in the form of forward guidance, discount rates, or borrowing constraints—because it intensifies the feedback loop between aggregate spending and inflation. By intensifying this kind of macroeconomic complementarity, however, a higher degree of price flexibility also raises the relative importance of higher-order beliefs, which in turn contributes to stronger attenuation effects of the type we have documented in this paper. In a nutshell, the very same mechanism that creates the paradox of flexibility within the New Keynesian framework also helps contain that paradox once we relax the common-knowledge assumption of that framework.

Note that we have proved the above result only under the restriction $\lambda_f = 1$, which means that only the consumers lack common knowledge. Whenever $\lambda_f < 1$, there is a conflicting effect, which is that higher price flexibility reduces the strategic complementarity that operates within the supply block, thereby also reducing the role of λ_f itself. For the numerical example considered earlier, however, the overall effect of higher price flexibility is qualitatively the same whether $\lambda_f = 1$ or $\lambda_f = \lambda_c$.

We illustrate this in Figure 2. We let $\lambda_f = \lambda_c = 0.75$, use the same parameter values as those used in Figure 1, and plot the relation between the ratio ϕ_T/ϕ_T^* and the horizon T under two values for θ. The solid red line corresponds to a higher value for θ, while the dashed blue line corresponds to a lower value for θ, that is, to more price flexibility. As evident in the figure, more price flexibility maps, not only to a lower ratio ϕ_T/ϕ_T^* (i.e., stronger attenuation) for any given T, but also to a more rapid decay in that ratio as we raise T.

Proof of Proposition 11. Consider first the environment studied in Section V and let us study the cross-partial derivative of the power of forward guidance with respect to κ and λ_c. To simplify notation, we use ϕ_τ and ϖ_τ as shortcuts for, respectively, $\phi\left(\lambda_c, \lambda_f, \tau\right)$ and $\varpi\left(\lambda_c, \lambda_f, \tau\right)$, where the functions ϕ and ϖ are defined as in the proof of Proposition 7. From conditions (66), we have

$$\frac{\partial \phi_1}{\partial \kappa} = \frac{\partial \phi(\lambda_c, 1, 1)}{\partial \kappa} = \sigma^2 > 0 \qquad \text{and} \qquad \frac{\partial \varpi_1}{\partial \kappa} = \frac{\partial \varpi(\lambda_c, 1, 1)}{\partial \kappa} = \sigma\beta > 0. \tag{114}$$

For any $\tau \geq 2$, when $\lambda_f = 1$, conditions (71) and (72) become

$$\phi_\tau = (\beta + (1 - \beta + \sigma\kappa)\,\lambda_c)\,\phi_{\tau-1} + \sigma\varpi_{\tau-1} \qquad \text{and} \qquad \varpi_\tau = \kappa\beta\lambda_c\phi_{\tau-1} + \beta\varpi_{\tau-1}.$$

As a result, for all $\tau \geq 2$, we have

$$\frac{\partial \phi_\tau}{\partial \kappa} = (\beta + (1 - \beta + \sigma\kappa)\,\lambda_c)\,\frac{\partial \phi_{\tau-1}}{\partial \kappa} + \sigma\lambda_c\phi_{\tau-1} + \sigma\frac{\partial \varpi_{\tau-1}}{\partial \kappa}, \tag{115}$$

$$\frac{\partial \varpi_\tau}{\partial \kappa} = \kappa\beta\lambda_c\frac{\partial \phi_{\tau-1}}{\partial \kappa} + \beta\lambda_c\phi_{\tau-1} + \beta\frac{\partial \varpi_{\tau-1}}{\partial \kappa}. \tag{116}$$

From conditions (114), (115) and (116), $\frac{\partial \phi_\tau}{\partial \kappa}$ and $\frac{\partial \varpi_\tau}{\partial \kappa}$ are strictly positive for any $\tau \geq 1$ by induction. Moreover, from conditions (66), (114), (115) and (116), we have that $\frac{\partial \phi_2}{\partial \kappa}$ and $\frac{\partial \varpi_2}{\partial \kappa}$ are strictly increasing in λ_c. Then, from conditions (115), (116) and the fact that ϕ_τ itself is strictly increasing in λ_c for all $\tau \geq 2$, we have $\frac{\partial \phi_\tau}{\partial \kappa}$ and $\frac{\partial \varpi_\tau}{\partial \kappa}$ are strictly increasing in λ_c for all $\tau \geq 2$ by induction.

Consider now the environment studied in Section VI and let us study the cross-partial derivative of the relevant fiscal multiplier with respect to κ and λ_c when $\lambda_f = 1$. From the proof of Proposition 8, similarly to conditions (114), (115) and (116), we have

$$\frac{\partial \phi_{g,1}}{\partial \kappa} = \sigma\,(1 - \Omega_c) > 0 \qquad \text{and} \qquad \frac{\partial \varpi_{g,1}}{\partial \kappa} = \beta\,(1 - \Omega_c) > 0,$$

$$\frac{\partial \phi_{g,\tau}}{\partial \kappa} = (\beta + (1 - \beta + \sigma\kappa\Omega_c)\,\lambda_c)\,\frac{\partial \phi_{g,\tau-1}}{\partial \kappa} + \sigma\Omega_c\lambda_c\phi_{g,\tau-1} + \sigma\frac{\partial \varpi_{g,\tau-1}}{\partial \kappa},$$

$$\frac{\partial \varpi_{g,\tau}}{\partial \kappa} = \kappa\beta\lambda_c\Omega_c\frac{\partial \phi_{g,\tau-1}}{\partial \kappa} + \beta\lambda_c\Omega_c\phi_{g,\tau-1} + \beta\frac{\partial \varpi_{g,\tau-1}}{\partial \kappa}.$$

The result then follows from the same argument as before.

Extension of Lemma 2 and Theorems 1 and 2. Here we show that Lemma 2, Theorem 2 and, by implication, Theorem 1 extend to the kind of multi-layer beauty contest seen in condition (22) of Lemma 3.

Similar to Section IV, we impose Assumption 1. Similar to the proof of Lemma 2, we can find positively-valued coefficients $\{\chi_{h,\tau}\}_{\tau\geq 1, 1\leq h\leq\tau}$, such that, for any $t \leq T-1$,

$$\tilde{y}_t - \tilde{y}_t^{trap} = \sum_{h=1}^{T-t}\left\{\chi_{h,T-t}\bar{E}_t^h\left[\tilde{R}_T\right]\right\}, \tag{117}$$

with $\tilde{y}_t^{trap}$ defined as in the proof of Proposition 6 and

$$\chi_{1,\tau} = \sigma \left(1 + \tau\sigma\kappa\right) \beta^{\tau-1} \quad \forall \tau \geq 1, \tag{118}$$

$$\chi_{k,\tau} = \sum_{l=1}^{\tau-k+1} \left(1 - \beta + l\sigma\kappa\right) \beta^{l-1} x_{k-1,\tau-l} \quad \forall k \geq 2 \; and \; \tau \geq k. \tag{119}$$

We can then characterize the combined effect of beliefs of order up to k on spending, $s_{k,\tau}$, as[20]

$$s_{k,\tau} = \sigma \left(1 + \tau\sigma\kappa\right) \beta^{\tau-1} + \sum_{l=1}^{\tau-1} \left(1 - \beta + l\sigma\kappa\right) \beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 1. \tag{120}$$

Let $d_\tau = s_{\tau,\tau}$ denote the combined effect of beliefs of all different orders on spending. Similar to condition (18), $d_\tau = \phi_\tau^*$. Following the proof of Proposition 6, we have

$$d_0 = \sigma \qquad \text{and} \qquad d_1 = \sigma \left(1 + \sigma\kappa\right),$$

$$\frac{d_\tau}{d_{\tau-1}} + \beta \frac{d_{\tau-2}}{d_{\tau-1}} = 1 + \beta + \sigma\kappa \quad \forall \tau \geq 2. \tag{121}$$

Now we prove $s_{k,\tau}$ satisfies an inequality with a similar form as condition (121):

$$\frac{s_{k,\tau}}{s_{k,\tau-1}} + \beta \frac{s_{k,\tau-2}}{s_{k,\tau-1}} \leq 1 + \beta + \sigma\kappa \quad \forall \tau \geq 3 \; and \; k \geq 1. \tag{122}$$

From condition (120), we have

$$\beta s_{k,\tau-1} = \sigma \left(1 + \left(\tau - 1\right) \sigma\kappa\right) \beta^{\tau-1} + \sum_{l=2}^{\tau-1} \left(1 - \beta + \left(l - 1\right) \sigma\kappa\right) \beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 2.$$

As a result, we have

$$s_{k,\tau} = \beta s_{k,\tau-1} + \left(1 - \beta\right) s_{k-1,\tau-1} + \sigma^2 \kappa \beta^{\tau-1} + \sigma\kappa \sum_{l=1}^{\tau-1} \beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 2,$$

$$\beta s_{k,\tau-1} = \beta^2 s_{k,\tau-2} + \beta \left(1 - \beta\right) s_{k-1,\tau-2} + \sigma^2 \kappa \beta^{\tau-1} + \sigma\kappa \sum_{l=2}^{\tau-1} \beta^{l-1} s_{k-1,\tau-l} \quad \forall k \geq 1 \; and \; \tau \geq 3.$$

Using the previous two conditions, we have, for all $k \geq 1$ and $\tau \geq 3$,

$$s_{k,\tau} + \beta^2 s_{k,\tau-2} + \beta \left(1 - \beta\right) s_{k-1,\tau-2} = 2\beta s_{k,\tau-1} + \left(1 - \beta + \sigma\kappa\right) s_{k-1,\tau-1},$$

$$s_{k,\tau} + \beta s_{k,\tau-2} = \left(1 + \beta + \sigma\kappa\right) s_{k,\tau-1} + \beta \left(1 - \beta\right) \chi_{k,\tau-2} - \left(1 - \beta + \sigma\kappa\right) \chi_{k,\tau-1}. \tag{123}$$

[20] Similar to the proof of Theorem 2, for notation simplicity, we extend the definition of $s_{h,\tau} = \sum_{r=1}^{h} \chi_{r,\tau}$ for all $h > \tau$. As for $h > \tau$, $\chi_{h,\tau} = 0$, we have $s_{h,\tau} = s_{\tau,\tau}$ for all $h > \tau$. We also define $s_{0,\tau} = 0$ for all $\tau \geq 1$.

To prove (122), we only need to prove:

$$\beta\left(1-\beta\right)\chi_{k,\tau-2}\leq\left(1-\beta+\sigma\kappa\right)x_{k,\tau-1}\quad\forall k\geq1\;and\;\tau\geq3. \tag{124}$$

In fact, we prove the following stronger result:

$$\beta\chi_{k,\tau-2}\leq x_{k,\tau-1}\quad\forall k\geq1\;and\;\tau\geq3. \tag{125}$$

From condition (118), we know that (125) is true for $k=1$ and $\tau\geq3$. From condition (119), we know that

$$\chi_{k,\tau-1}=\sum_{l=1}^{\tau-k}\left(1-\beta+\sigma l\kappa\right)\beta^{l-1}x_{k-1,\tau-1-l}\quad\forall k\geq2\;and\;\tau\geq k+1, \tag{126}$$

$$\beta\chi_{k,\tau-2}=\sum_{l=2}^{\tau-k}\left(1-\beta+\sigma\left(l-1\right)\kappa\right)\beta^{l-1}x_{k-1,\tau-1-l}\quad\forall k\geq2\;and\;\tau\geq k+2.$$

This proves $\beta\chi_{k,\tau-2}\leq x_{k,\tau-1}$ for $k\geq2$ and $\tau\geq k+2$. Together with the fact that, $\chi_{k,\tau-2}=0\;\forall k\geq\tau-1$, we prove (125) and thus (124). This finishes the proof of (122).

Based on (121) and (122), we can then establish a result akin to Theorem 2. That is, for any given $k\geq1$ and $\tau\geq k$, the relative contribution of the first k orders, $\frac{s_{k,\tau}}{s_{\tau,\tau}}=\frac{s_{k,\tau}}{d_\tau}$, strictly decreases with τ.

First, note that, for any given $k\geq1$,$1=\frac{s_{k,k}}{d_k}>\frac{s_{k,k+1}}{d_{k+1}}$, because $x_{k+1,k+1}>0$. Then, we can proceed by induction on $\tau\geq k$, for any fixed $k\geq1$. If we have $\frac{s_{k,\tau}}{d_\tau}>\frac{s_{k,\tau+1}}{d_{\tau+1}}$ for some $\tau\geq k$, we have $\frac{s_{k,\tau}}{s_{k,\tau+1}}>\frac{d_\tau}{d_{\tau+1}}$. Using (121) and (122), we have $\frac{s_{k,\tau+2}}{s_{k,\tau+1}}<\frac{d_{\tau+2}}{d_{\tau+1}}$, and thus $\frac{s_{k,\tau+1}}{d_{\tau+1}}>\frac{s_{k,\tau+2}}{d_{\tau+2}}$. This completes the proof that, for any $k\geq1$ and any $\tau\geq k$, the ratio $\frac{s_{k,\tau}}{s_{\tau,\tau}}$, strictly decreases with the horizon τ.

Finally, we prove that, for any $k\geq1$,

$$\frac{s_{k,\tau}}{s_{\tau,\tau}}\to0,\quad as\;\tau\to\infty. \tag{127}$$

In other words, we want to prove the relative contribution of the first k orders of beliefs to aggregate spending converges to zero when the horizon τ goes to infinity.

First note that, from condition (120), we have $s_{1,\tau}=\sigma\left(1+\sigma\tau\kappa\right)\beta^{\tau-1}\to0,$ as $\tau\to+\infty$. From the proof of Proposition 6, we know $s_{\tau,\tau}=d_\tau=\phi^*_\tau\geq\sigma$. As a result, (127) is true for $k=1$.

If there exists $k\geq2$ such that (127) does not hold, we let $k^*\geq2$ denote the smallest of such k. Then, (127) holds for $1\leq k\leq k^*-1$. Because we already prove that $\frac{s_{k^*,\tau}}{s_{\tau,\tau}}\geq0$ is decreasing with the horizon τ, there exists $0<\Gamma<1$ such that $\frac{s_{k^*,\tau}}{s_{\tau,\tau}}=\frac{s_{k^*,\tau}}{\phi^*_\tau}\to\Gamma$ as $\tau\to\infty$. As a result, $\frac{s_{k^*,\tau}}{\phi^*_\tau}\frac{\phi^*_{\tau-1}}{s_{k^*,\tau-1}}\to1$ as $\tau\to\infty$. Because we already prove that, in the proof of Proposition 6, $\frac{\phi^*_\tau}{\phi^*_{\tau-1}}\to\Gamma^*$, we have

$$\frac{s_{k^*,\tau}}{s_{k^*,\tau-1}}\to\Gamma^*\quad and\quad\frac{s_{k^*,\tau-2}}{s_{k^*,\tau-1}}\to\frac{1}{\Gamma^*}\quad as\;\tau\to\infty. \tag{128}$$

Note that since $s_{k^*,\tau}=s_{k^*-1,\tau}+\chi_{k^*,\tau}$ and $\frac{s_{k^*-1,\tau}}{s_{\tau,\tau}}\to0\;as\;\tau\to\infty$, we have $\frac{\chi_{k^*,\tau}}{s_{\tau,\tau}}=\frac{\chi_{k^*,\tau}}{\phi^*_\tau}\to\Gamma\;as\;\tau\to\infty$.

As a result,

$$\frac{\chi_{k^*,\tau}}{s_{k^*,\tau}} = \frac{\chi_{k^*,\tau}}{\phi^*_\tau}\frac{\phi^*_\tau}{s_{k^*,\tau}} \to 1 \quad as\ \tau \to \infty. \tag{129}$$

Now we prove a stronger version of (122)

$$\frac{s_{k,\tau}}{s_{k,\tau-1}} + \beta\frac{s_{k,\tau-2}}{s_{k,\tau-1}} + \sigma\kappa\frac{\chi_{k,\tau-1}}{s_{k,\tau-1}} \leq 1+\beta+\sigma\kappa \quad \forall\tau \geq 3\ and\ k \geq 1. \tag{130}$$

This comes from the fact that (125) can be written as

$$\beta\left(1-\beta\right)\chi_{k,\tau-2} + \sigma\kappa\chi_{k,\tau-1} \leq \left(1-\beta+\sigma\kappa\right)x_{k,\tau-1} \quad \forall\tau \geq 3\ and\ k \geq 1. \tag{131}$$

Using (61), (128) and (129), we have

$$\frac{s_{k^*,\tau}}{s_{k^*,\tau-1}} + \beta\frac{s_{k^*,\tau-2}}{s_{k^*,\tau-1}} + \sigma\kappa\frac{\chi_{k^*,\tau-1}}{s_{k^*,\tau-1}} \to \Gamma^* + \beta\frac{1}{\Gamma^*} + \sigma\kappa = 1+\beta+2\sigma\kappa \quad as\ \tau \to \infty.$$

This contradicts (130) when $\kappa > 0$ and proves (127). This finishes the proof of the result akin to Theorem 2. Together with Proposition 3, we then establish the "horizon effect" akin to Theorem 1. Similarly, Proposition 4 also holds here.

References

Angeletos, George-Marios, and Chen Lian. (2016) "Forward Guidance without Common Knowledge." *NBER Working Paper No. 22785*.

Coibion, Olivier, and Yuriy Gorodnichenko. (2012) "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy*, 120(1): 116–159.

Eggertsson, Gauti B, and Paul Krugman. (2012) "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach." *The Quarterly Journal of Economics*, 127(3): 1469–1513.

Gabaix, Xavier. (2016) "A Behavioral New Keynesian Model." *NBER Working Paper No. 22954*.

Khaw, Mel Win, Luminita Stevens, and Michael Woodford. (2016) "Discrete Adjustment to a Changing Environment: Experimental Evidence." *NBER Working Paper No. 22978*.

Luo, Yulei, Jun Nie, Gaowang Wang, and Eric R. Young. (2017) "Rational inattention and the dynamics of consumption and wealth in general equilibrium." *Journal of Economic Theory*, 172(C): 55 – 87.

Mackowiak, Bartosz, and Mirko Wiederholt. (2009) "Optimal Sticky Prices under Rational Inattention." *The American Economic Review*, 99(3): 769–803.

Mackowiak, Bartosz, Filip Matejka, and Mirko Wiederholt. (2017) "The Rational Inattention Filter." *Goethe University Frankfurt mimeo*.

Mankiw, Gregory, and Ricardo Reis. (2002) “Sticky Information Versus Sticky Prices: a Proposal to Replace the New Keynesian Phillips Curve.” *The Quarterly Journal of Economics*, 117(4): 1295–1328.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. (2016) “The Power of Forward Guidance Revisited.” *The American Economic Review*, 106(10): 3133–3158.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. (2017) “The Discounted Euler Equation: A Note.” *Economica*, 84(336): 820–831.

Nimark, Kristoffer. (2008) “Dynamic Pricing and Imperfect Common Knowledge.” *Journal of Monetary Economics*, 55(2): 365–382.

Tirole, Jean. (2015) “Cognitive Games and Cognitive Traps.” *Toulouse School of Economics mimeo*.

Werning, Iván. (2015) “Incomplete Markets and Aggregate Demand.” *NBER Working Paper No. 21448*.

Wiederholt, Mirko. (2015) “Empirical Properties of Inflation Expectations and the Zero Lower Bound.” *Goethe University Frankfurt mimeo*.

Woodford, Michael. (2003) “Imperfect Common Knowledge and the Effects of Monetary Policy.” *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*.

Woodford, Michael. (2011) “Simple Analytics of the Government Expenditure Multiplier.” *American Economic Journal: Macroeconomics*, 3(1): 1–35.