Online Appendix to "Rational Inattention and the Business Cycle Effects of Productivity and News Shocks"

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
LMU Munich, Sciences Po, and CEPR

October 2023

Abstract

This document contains Appendices A-F for the October 2023 version of the paper "Rational Inattention and the Business Cycle Effects of Productivity and News Shocks."

A Non-stochastic steady state

The non-stochastic steady state is the solution of the model when total factor productivity e^{a_t} is equal to 1 in every period, this is common knowledge, and all variables are constant over time.

Let an upper-case letter without a time subscript denote the value of a variable in the non-stochastic steady state. Profit maximization implies that $\alpha K_i^{\alpha-1} L_i^{\phi} N_i^{1-\alpha-\phi} = \beta^{-1} - 1 + \delta$ and $\phi K_i^{\alpha} L_i^{\phi-1} N_i^{1-\alpha-\phi} = W$ for each firm i, which determines K_i and L_i as functions of W and parameter values (including N_i):

$$K_i = \left(\frac{\alpha}{\beta^{-1} - 1 + \delta}\right)^{\frac{1-\phi}{1-\alpha-\phi}} \left(\frac{\phi}{W}\right)^{\frac{\phi}{1-\alpha-\phi}} N_i$$

$$L_i = \left(\frac{\alpha}{\beta^{-1} - 1 + \delta}\right)^{\frac{\alpha}{1-\alpha-\phi}} \left(\frac{\phi}{W}\right)^{\frac{1-\alpha}{1-\alpha-\phi}} N_i.$$

Suppose that N_i is constant across i, $N_i = N$. It follows that K_i and L_i are constant across i, $K_i = K$, $L_i = L$. Moreover, Y_i, I_i and D_i are also constant across i, $Y_i = Y = K^{\alpha} L^{\phi} N^{1-\alpha-\phi}$, $I_i = I = \delta K$, $D_i = D = Y - WL - I$.

Utility maximization implies that $V = [\beta/(1-\beta)]\, D$ and $W C_j^{-\gamma} = L_j^{\eta}$ for each household j. Suppose that in the non-stochastic steady state each household holds an equal share of the mutual fund, $Q_j = 1$ for each j. C_j and L_j are then constant across j, $C_j = C$, $L_j = L$, and the budget constraint implies that $C = WL + D$. Combining this equation with $D = Y - WL - I$ yields the resource constraint $Y = C + I$.

One can solve the system of equations:

$$K = \left(\frac{\alpha}{\beta^{-1} - 1 + \delta}\right)^{\frac{1-\phi}{1-\alpha-\phi}} \left(\frac{\phi}{W}\right)^{\frac{\phi}{1-\alpha-\phi}} N$$

$$L = \left(\frac{\alpha}{\beta^{-1} - 1 + \delta}\right)^{\frac{\alpha}{1-\alpha-\phi}} \left(\frac{\phi}{W}\right)^{\frac{1-\alpha}{1-\alpha-\phi}} N$$

$$W = L^{\eta} \left(K^{\alpha} L^{\phi} N^{1-\alpha-\phi} - \delta K\right)^{\gamma}$$

for K, L and W for given parameter values (including N). The last equation comes from combining the equilibrium condition $WC^{-\gamma} = L^{\eta}$ with the resource constraint. One can then compute the other endogenous variables (Y, I, C, D, and V) from the equations $Y = K^{\alpha} L^{\phi} N^{1-\alpha-\phi}$, $I = \delta K$, $C = Y - I$, $D = Y - WL - I$, $V = [\beta/(1-\beta)]\, D$.

The following steady-state ratios are useful: $WL/Y = \phi$, $I/Y = \alpha\beta\delta/[1 - \beta(1-\delta)]$, $C/Y = 1 - I/Y$, $D/Y = 1 - WL/Y - I/Y$, $V/C = [\beta/(1-\beta)]\,(D/Y)(Y/C)$ (WL/C and D/C follow).

B Perfect information benchmark

Suppose that all agents have perfect information. Let a lower-case letter denote the log-deviation of a variable from its value in the non-stochastic steady state. The firms' first-order conditions imply that

$$a_t + \alpha k_{t-1} - (1-\phi)\, l_t = w_t$$

$$E_t a_{t+1} - (1-\alpha)\, k_t + \phi E_t l_{t+1} = \frac{\gamma\,(E_t c_{t+1} - c_t)}{1 - \beta\,(1-\delta)}.$$

From the production function, the law of motion of capital and the profit function, we have

$$y_t = a_t + \alpha k_{t-1} + \phi l_t$$

$$\delta i_t = k_t - (1-\delta)\, k_{t-1}$$

$$(D/Y)\, d_t = y_t - (WL/Y)\,(w_t + l_t) - (I/Y)\, i_t.$$

The households' first-order conditions imply that

$$\gamma E_t \left(c_{t+1} - c_t\right) = \beta E_t v_{t+1} - v_t + \left(1 - \beta\right) E_t d_{t+1}$$

$$w_t - \gamma c_t = \eta l_t.$$

The resource constraint reads

$$y_t = \left(C/Y\right) c_t + \left(I/Y\right) i_t.$$

C Expected loss in profit from suboptimal actions

Proposition 1 *Let $E_{i,-1}$ denote the expectation operator conditioned on information of the decision-maker of firm i in period -1. Let g denote the functional that is obtained by multiplying the profit function by β^t and summing over all t from zero to infinity. Let $\tilde{g}$ denote the second-order Taylor approximation of g at the non-stochastic steady state. Let χ_t, z_t and v_t denote the following vectors*

$$\chi_t = \begin{pmatrix} k_{it} \\ l_{it} \end{pmatrix} \quad z_t = \begin{pmatrix} a_t \\ w_t \\ c_t \end{pmatrix} \quad v_t = \begin{pmatrix} \chi_t \\ z_t \\ 1 \end{pmatrix}.$$

Suppose that the decision-maker of firm i knows in period -1 the firm's initial capital stock, $k_{i,-1}$. Suppose also that there exist two constants $\delta < (1/\beta)$ and $A \in \mathbb{R}$ such that, for each period $t \geq 0$, for $\tau = 0, 1$, and for all $m, n \in \{1, 2, 3, 4, 5, 6\}$,

$$E_{i,-1} \left| v_{m,t} v_{n,t+\tau} \right| < \delta^t A. \tag{1}$$

Here $v_{m,t}$ and $v_{n,t}$ denote the mth and nth element of the vector v_t. Then the expected discounted sum of losses in profit when the law of motion for the actions differs from the law of motion for the optimal actions under perfect information equals

$$\begin{aligned} & E_{i,-1}\left[\tilde{g}\left(k_{i,-1}, \chi_0, z_0, \chi_1, z_1, \ldots\right)\right] - E_{i,-1}\left[\tilde{g}\left(k_{i,-1}, \chi_0^*, z_0, \chi_1^*, z_1, \ldots\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left[\frac{1}{2}\left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right)\right]. \end{aligned} \tag{2}$$

The matrices Θ_0 and Θ_1 are given by

$$\Theta_0 = -C^{-\gamma} Y \begin{bmatrix} \beta\alpha\left(1-\alpha\right) & 0 \\ 0 & \phi\left(1-\phi\right) \end{bmatrix} \quad \Theta_1 = C^{-\gamma} Y \begin{bmatrix} 0 & \beta\alpha\phi \\ 0 & 0 \end{bmatrix}. \tag{3}$$

The optimal actions under perfect information are given by

$$\chi_t^* = \begin{pmatrix} k_{it}^* \\ l_{it}^* \end{pmatrix} = \begin{pmatrix} \frac{1}{1-\alpha-\phi}\left[E_t a_{t+1} - \phi E_t w_{t+1} - (1-\phi)\frac{\gamma E_t(c_{t+1}-c_t)}{1-\beta(1-\delta)}\right] \\ \frac{\alpha}{1-\phi} k_{it-1}^* + \frac{1}{1-\phi}(a_t - w_t) \end{pmatrix}, \tag{4}$$

where E_t denotes the expectation operator conditioned on the entire history up to and including period t, and the initial capital stock is given by the initial condition $k_{i,-1}^* = k_{i,-1}$.

Proof: First, we introduce notation. The profit of firm i in period t depends on three sets of variables: (i) variables that the decision-maker of firm i chooses in period t (here k_{it} and l_{it}), (ii) variables that the decision-maker chose in the past (here k_{it-1}), and (iii) variables that the decision-maker takes as given (here a_t, w_t and c_t). The first set of variables is collected in the vector χ_t, the second set of variables is an element of χ_{t-1} for all $t \geq 0$ once we define the vector $\chi_{-1} = (k_{i,-1}, 0)'$, and the third set of variables is collected in the vector z_t.

The next steps are word for word identical to steps "Second" to "Seventh" in proof of Proposition 2 in online Appendix D of Maćkowiak and Wiederholt (2015). The reason is that these steps only require that the payoff in period t depends only on own current actions (collected in χ_t), own previous-period actions (collected in χ_{t-1}), and variables that the decision-maker takes as given (collected in z_t) and that the initial condition $k_{i,-1}$ and the vector v_t satisfy conditions (40)-(42) in online Appendix D of Maćkowiak and Wiederholt (2015). The payoff in period t in Proposition 1 is profit, whereas the payoff in period t in online Appendix D of Maćkowiak and Wiederholt (2015) is period utility, but in both cases this payoff depends only on own current actions (χ_t), own previous-period actions (χ_{t-1}), and variables that the decision-maker takes as given (z_t). Conditions (40)-(41) in online Appendix D of Maćkowiak and Wiederholt (2015) are satisfied because of the assumption in Proposition 1 that the decision-maker knows the initial condition χ_{-1}. Condition (42) in online Appendix D of Maćkowiak and Wiederholt (2015) is equal to condition (1) in Proposition 1. These steps "Second" to "Seventh" yield equation (2), where Θ_0 is defined as the Hessian matrix of second derivatives of g with respect to χ_t evaluated at the non-stochastic steady state and divided by β^t, Θ_1 is defined as the Hessian matrix of second derivatives of g with respect to χ_t and χ_{t+1} evaluated at the non-stochastic steady state and divided by β^t, and χ_t^* is defined as the actions that the decision-maker would take if he or she had perfect information in every period $t \geq 0$.

Eighth, the functional g in Proposition 1 is the discounted sum of profit

$$g\left(\chi_{-1},\chi_0,z_0,\chi_1,z_1,\ldots\right)=\sum_{t=0}^{\infty}\beta^t f\left(\chi_t,\chi_{t-1},z_t\right),$$

where the function f is the profit function

$$f\left(\chi_t,\chi_{t-1},z_t\right)=C^{-\gamma}e^{-\gamma c_t}Y\left\{e^{a_t+\alpha k_{it-1}+\phi l_{it}}-\phi e^{w_t+l_{it}}+\left(\frac{\alpha}{\beta^{-1}-1+\delta}\right)\left[\left(1-\delta\right)e^{k_{it-1}}-e^{k_{it}}\right]\right\}.$$

Computing the matrices Θ_0 and Θ_1 for this functional g yields equation (3).

Ninth, we characterize the optimal actions under perfect information. Formally, the process $\{\chi_t^*\}_{t=0}^{\infty}$ is defined by the initial condition $\chi_{-1}^*=(k_{i,-1},0)'$ and the optimality condition

$$\forall t\geq 0: E_t\left[\theta_0+\Theta_{-1}\chi_{t-1}^*+\Theta_0\chi_t^*+\Theta_1\chi_{t+1}^*+\Phi_0 z_t+\Phi_1 z_{t+1}\right]=0. \tag{5}$$

Here θ_0 is defined as the vector of first derivatives of g with respect to χ_t evaluated at the non-stochastic steady state and divided by β^t, Θ_{-1} is defined as the matrix of second derivatives of g with respect to χ_t and χ_{t-1} evaluated at the non-stochastic steady state and divided by β^t, Φ_0 is defined as the matrix of second derivatives of g with respect to χ_t and z_t evaluated at the non-stochastic steady state and divided by β^t, Φ_1 is defined as the matrix of second derivatives of g with respect to χ_t and z_{t+1} evaluated at the non-stochastic steady state and divided by β^t, and E_t denotes the expectation operator conditioned on the entire history up to and including period t. See the step "Fourth" in proof of Proposition 2 in online Appendix D of Maćkowiak and Wiederholt (2015). Computing the vector θ_0 and the matrices Θ_{-1}, Φ_0, and Φ_1 for the functional g defined in the previous step yields

$$\begin{array}{cc} \theta_0=\begin{pmatrix}0\\0\end{pmatrix} & \Theta_{-1}=C^{-\gamma}Y\begin{bmatrix}0 & 0\\ \alpha\phi & 0\end{bmatrix} \\ \Phi_0=C^{-\gamma}Y\begin{bmatrix}0 & 0 & \gamma\frac{\alpha}{\beta^{-1}-1+\delta}\\ \phi & -\phi & 0\end{bmatrix} & \Phi_1=C^{-\gamma}Y\begin{bmatrix}\alpha\beta & 0 & -\gamma\frac{\alpha}{\beta^{-1}-1+\delta}\\ 0 & 0 & 0\end{bmatrix}\end{array}.$$

Substituting the equations for θ_0, Θ_{-1}, Φ_0, Φ_1, Θ_0, Θ_1, χ_t and z_t into equation (5) yields

$$E_t a_{t+1}-\left(1-\alpha\right)k_{it}^*+\phi E_t l_{it+1}^*=\frac{\gamma E_t\left(c_{t+1}-c_t\right)}{1-\beta\left(1-\delta\right)} \tag{6}$$

$$a_t+\alpha k_{it-1}^*-\left(1-\phi\right)l_{it}^*=w_t. \tag{7}$$

Equations (6)-(7) are the usual optimality conditions for capital and labor. Equation (6) states that the profit-maximizing capital input equates the expected marginal product of capital to the cost of capital. Equation (7) states that the profit-maximizing labor input equates the marginal product of labor to the wage. Rearranging equations (6)-(7) yields the closed-form solution (4) for the actions that the firm would take in period t if the firm had perfect information in every period $t \geq 0$.

Proposition 2 *Under condition (1), we have*

$$\begin{aligned} & \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{1}{2} \left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right) \right] \\ = & -C^{-\gamma} Y \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left[\frac{\beta\alpha \left(1 - \alpha - \frac{\alpha\phi}{1-\phi}\right)}{2} \left(k_{it} - k_{it}^*\right)^2 + \frac{\phi\left(1-\phi\right)}{2} \left(\zeta_{it} - \zeta_{it}^*\right)^2 \right], \end{aligned} \tag{8}$$

where $\zeta_{it} \equiv l_{it} - \frac{\alpha}{1-\phi} k_{it-1}$ *and* $\zeta_{it}^* \equiv l_{it}^* - \frac{\alpha}{1-\phi} k_{it-1}^*$.

Proof: First, we have

$$\begin{aligned} & \frac{1}{2} \left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right) \\ = & C^{-\gamma} Y \left[-\frac{\beta\alpha\left(1-\alpha\right)}{2} \left(k_{it} - k_{it}^*\right)^2 - \frac{\phi\left(1-\phi\right)}{2} \left(l_{it} - l_{it}^*\right)^2 + \beta\alpha\phi \left(k_{it} - k_{it}^*\right) \left(l_{it+1} - l_{it+1}^*\right) \right] \end{aligned} \tag{9}$$

because $\chi_t = (k_{it}, l_{it})'$, $\chi_t^* = (k_{it}^*, l_{it}^*)'$ and the matrices Θ_0 and Θ_1 are given by equation (3). Second, define

$$\zeta_{it} = l_{it} - \frac{\alpha}{1-\phi} k_{it-1},$$

and

$$\zeta_{it}^* = l_{it}^* - \frac{\alpha}{1-\phi} k_{it-1}^*.$$

This definition implies

$$l_{it} - l_{it}^* = \zeta_{it} - \zeta_{it}^* + \frac{\alpha}{1-\phi} \left(k_{it-1} - k_{it-1}^*\right).$$

Substituting the last equation into equation (9) yields

$$\begin{aligned} & \frac{1}{2} \left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right) \\ = & C^{-\gamma} Y \left[\begin{array}{c} -\frac{\beta\alpha(1-\alpha)}{2} \left(k_{it} - k_{it}^*\right)^2 - \frac{\phi(1-\phi)}{2} \left(\zeta_{it} - \zeta_{it}^*\right)^2 \\ -\alpha\phi \left(\zeta_{it} - \zeta_{it}^*\right) \left(k_{it-1} - k_{it-1}^*\right) - \frac{\alpha^2\phi}{2(1-\phi)} \left(k_{it-1} - k_{it-1}^*\right)^2 \\ +\beta\alpha\phi \left(k_{it} - k_{it}^*\right) \left(\zeta_{it+1} - \zeta_{it+1}^*\right) + \frac{\beta\alpha^2\phi}{1-\phi} \left(k_{it} - k_{it}^*\right)^2 \end{array} \right]. \end{aligned}$$

Multiplying by β^t and summing over all t from zero to T yields

$$\begin{aligned}
&\sum_{t=0}^{T}\beta^t\left[\frac{1}{2}\left(\chi_t-\chi_t^*\right)'\Theta_0\left(\chi_t-\chi_t^*\right)+\left(\chi_t-\chi_t^*\right)'\Theta_1\left(\chi_{t+1}-\chi_{t+1}^*\right)\right]\\
=\;& C^{-\gamma}Y\left[\begin{array}{c}-\frac{\beta\alpha}{2}\left(1-\alpha-\frac{\alpha\phi}{1-\phi}\right)\sum_{t=0}^{T}\beta^t\left(k_{it}-k_{it}^*\right)^2-\frac{\phi(1-\phi)}{2}\sum_{t=0}^{T}\beta^t\left(\zeta_{it}-\zeta_{it}^*\right)^2\\ +\beta^T\frac{\beta\alpha^2\phi}{2(1-\phi)}\left(k_{iT}-k_{iT}^*\right)^2+\beta\alpha\phi\beta^T\left(k_{iT}-k_{iT}^*\right)\left(\zeta_{iT+1}-\zeta_{iT+1}^*\right)\end{array}\right],
\end{aligned}$$

where we have used $k_{i,-1}^*=k_{i,-1}$ and the fact that several terms on the right-hand side cancel. Taking the expectation $E_{i,-1}$ and the limit as $T\to\infty$ yields

$$\begin{aligned}
&\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left[\frac{1}{2}\left(\chi_t-\chi_t^*\right)'\Theta_0\left(\chi_t-\chi_t^*\right)+\left(\chi_t-\chi_t^*\right)'\Theta_1\left(\chi_{t+1}-\chi_{t+1}^*\right)\right]\\
=\;& C^{-\gamma}Y\left[\begin{array}{c}-\frac{\beta\alpha}{2}\left(1-\alpha-\frac{\alpha\phi}{1-\phi}\right)\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left(k_{it}-k_{it}^*\right)^2-\frac{\phi(1-\phi)}{2}\sum_{t=0}^{\infty}\beta^t E_{i,-1}\left(\zeta_{it}-\zeta_{it}^*\right)^2\\ +\frac{\beta\alpha^2\phi}{2(1-\phi)}\lim_{T\to\infty}\beta^T E_{i,-1}\left(k_{iT}-k_{iT}^*\right)^2+\beta\alpha\phi\lim_{T\to\infty}\beta^T E_{i,-1}\left(k_{iT}-k_{iT}^*\right)\left(\zeta_{iT+1}-\zeta_{iT+1}^*\right)\end{array}\right].
\end{aligned}$$

Under condition (1) the two infinite sums on the right-hand side of the last equation converge to an element in $\mathbb{R}$ and the third and fourth term on the right-hand side of the last equation equal zero.

D Additional numerical results for Section 4

We report here additional numerical results that help develop intuition about optimal signals in an economy with news shocks. We consider the version of the model from Section 4.1 in the paper with labor as the only variable input, $\alpha=0$. All parameters have the same values as in Section 4.1 unless otherwise indicated. We solved for the rational inattention equilibrium in Section 4.1. Here we suppose that a measure zero of firms are subject to rational inattention and other firms (and all households) have perfect information. We study the attention problem of the rationally inattentive firms.

In the first experiment, we vary h in the law of motion for productivity, $a_t=\rho a_{t-1}+\sigma\varepsilon_{t-h}$, between $h=0$ and $h=6$. Appendix Table 1 reports how the optimal signal changes. With $h=0$ the optimal signal is on the current optimal action l_{it}^*, $S_{it}=l_{it}^*+\psi_{it}$, where ψ_{it} follows a Gaussian white noise process with standard deviation $\sigma_\psi=0.0191$. With $h\geq 1$ the optimal signal has non-zero weights on $\varepsilon_t,...,\varepsilon_{t-(h-1)}$ because all elements of the state vector $\xi_t=\left(l_{it}^*,\varepsilon_t,...,\varepsilon_{t-(h-1)}\right)'$ help predict future optimal actions. For example, with $h = 2$ the optimal weight on ε_t is 0.0045 and the optimal weight on ε_{t-1} is 0.0059 (with the weight on l^*_{it} normalized to one). The largest weight is always on $\varepsilon_{t-(h-1)}$, the innovation useful for predicting the optimal action in period $t+1, t+2, \ldots.$ The weights decline monotonically from $\varepsilon_{t-(h-1)}$ to ε_t, the innovation useful for predicting the optimal action in period $t+h, t+h+1, \ldots.$ For a given marginal cost λ, the chosen amount of attention falls with h (because the marginal benefit of attention falls with h). Appendix Table 1 also reports the impulse response of labor input on impact as a fraction of the maximum response under perfect information. The impulse response on impact decreases with h, approaching zero as h rises. The more distant is the change in productivity, the weaker is the response of the action on impact of a news shock.

In the second experiment, we set $h = 1$ ($a_t = \rho a_{t-1} + \sigma\varepsilon_{t-1}$) and we vary ρ, adjusting σ to keep constant the unconditional variance of the optimal action l^*_{it} (as ρ rises the optimal action becomes more persistent while its unconditional variance remains unchanged). The state vector is $\xi_t = (l^*_{it}, \varepsilon_t)'$. The optimal signal has a non-zero weight on ε_t, and we find that the weight on ε_t falls with ρ. See Appendix Table 2. With a more persistent productivity process, learning about the innovation ε_t becomes less important relative to learning about the current state of productivity a_t. In addition, (i) the expected profit loss declines with ρ, and (ii) there is a non-monotonic relation between ρ and the chosen amount of attention (as ρ falls the quality of tracking deteriorates, but the marginal value of attention may go up or down). See also Proposition 3 in Maćkowiak and Wiederholt (2009).

In the third experiment, we add another MA term in the law of motion for productivity. As an example, we suppose that productivity follows the process $a_{t+1} = \rho a_t + \pi\varepsilon_{t-1} + \sigma\varepsilon_{t-3}$. Information about productivity becomes available $h = 4$ periods in advance and, if $\pi \neq 0$, additional information becomes available in an intermediate period (two periods in advance). Appendix Table 3 shows what happens to the optimal signal as we vary π (we adjust σ to keep constant the unconditional variance of the optimal action). The state vector is $\xi_t = (l^*_{it}, \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3})'$. As π rises, it is optimal to increase the weights on ε_t and ε_{t-1} and to decrease the weights on ε_{t-2} and ε_{t-3} in the signal. Appendix Table 3 also reports the response of labor input on impact of a news shock as a fraction of the maximum response under perfect information. This ratio rises with π, indicating that the action becomes *more* front-loaded.

In the fourth experiment, we assume that productivity is driven by two orthogonal shocks, a standard productivity shock and a news shock: $a_t = \rho a_{t-1} + \sigma_1 \varepsilon_{1t} + \sigma_2 \varepsilon_{2,t-h}$, where ε_{1t} and ε_{2t} follow independent Gaussian white noise process with unit variance, $\sigma_1 > 0$, $\sigma_2 > 0$, and $h \geq 1$. The profit-maximizing labor input l^*_{it} is still proportional to a_t (equation (20) in the paper). Focus on $h = 1$. The optimal signal is a one-dimensional signal about the vector $(l^*_{it}, \varepsilon_{2t})'$, or equivalently $(a_t, \varepsilon_{2t})'$.[1] The only difference to the case when productivity follows the process $a_t = \rho a_{t-1} + \sigma \varepsilon_{t-1}$ is that the signal is on $(a_t, \varepsilon_{2t})'$ instead of $(a_t, \varepsilon_t)'$. To begin, let us split the variance of productivity *equally* between the component driven by the first shock (ε_{1t}) and the component driven by the second shock (ε_{2t}), and solve the attention problem. We find that labor input l_{it} rises on impact of a positive news shock. The impulse response of l_{it} to ε_{2t} (Appendix Figure 1, middle-left panel) is very similar to the impulse response of l_{it} to ε_t in the case with $a_t = \rho a_{t-1} + \sigma \varepsilon_{t-1}$ (Appendix Figure 1, middle-right panel, reproduced from the lower-left panel of Figure 1 in the paper). The response of l_{it} on impact of a news shock equals 0.16 as a fraction of the maximum response under perfect information; the same ratio equals 0.15 in the case when productivity follows the process $a_t = \rho a_{t-1} + \sigma \varepsilon_{t-1}$.

We can also split the variance of productivity *unequally* between the two shocks. The key qualitative finding is unchanged: l_{it} rises on impact of a positive news shock. The quantitative findings (the impulse response of l_{it} to ε_{2t}, and the impact ratio) change little, whether we split the variance 90-10 in favor of the news shock or 10-90 against the news shock. The impact ratio falls slightly when the relative variance of the news shock rises – the impact ratio is 0.15 in the 90-10 case and 0.17 in the 10-90 case. Since we assume a ρ close to 1, the incentive to learn about a_t dominates the incentive to learn about the news shock and therefore the optimal signal changes little with the relative variance of the two shocks.

So far we assumed that, after choosing the signal process in period -1, the agent receives a sequence of signals in period -1 such that the conditional second moments are independent of time. See the discussion of problem (10)-(16) in Section 3.2 in the paper. In the fifth experiment, we resolve problem (10)-(16) having dropped this assumption. We use the methodology of Afrouzi and

[1] For any $h \geq 1$ we can write $l^*_{it} = l^*_{i1t} + l^*_{i2t}$, where l^*_{i1t} follows an AR(1) process driven by the standard productivity shock and l^*_{i2t} follows an ARMA(1,h) process driven by the news shock. The optimal signal is a one-dimensional signal about the state vector with equal weights on l^*_{i1t} and l^*_{i2t}. See also Maćkowiak, Matějka, and Wiederholt (2018), Section 4.4.

Yang (2021) to compute the steady-state conditional variance of ξ_t given $\mathcal{I}_{it-1}$, the steady-state conditional variance of ξ_t given $\mathcal{I}_{it}$, and the implied action. This approach allows for transitional dynamics in the conditional second moments. As before in this appendix, we study the version of the model from Section 4.1 (we also continue to suppose that only a measure zero of firms are subject to rational inattention). We compare: (i) the impulse response of labor input to a news shock with the assumption about the initial sequence of signals made earlier, and (ii) the analogous impulse response without this assumption. The two impulse responses are *identical* in the limit as the discount factor β approaches 1. The question is by how much the two impulse responses differ in this model for the value of the discount factor we have assumed, $\beta = 0.99$. The lower panel in Appendix Figure 1 shows the comparison ($h = 1$ on the left and $h = 4$ on the right; in each case, the line with circles is the baseline and the line with asterisks is the solution allowing for the transitional dynamics). The two impulse responses are *almost identical*; the agent who solves the attention problem allowing for the transitional dynamics chooses to process slightly less information in the steady state. Thus, the two solutions differ very little quantitatively and the qualitative result is unchanged: Labor input rises on impact of a positive news shock.

E Additional numerical results for Section 5

We report additional numerical results for the rational inattention equilibrium from Section 5.

Shimer (2009) emphasizes that the labor wedge is countercyclical in the data. In this model, the labor wedge equals $(y_t - l_t) - (\gamma c_t + \eta l_t)$, i.e., the gap between the marginal product of labor and the marginal rate of substitution between consumption and leisure, in the aggregate economy. Conditional on a positive news shock ($h \geq 1$), in the RI equilibrium the labor wedge is negative so long as productivity remains unchanged – the labor wedge is countercyclical (the reason is that output and employment rise but consumption rises more strongly). Once productivity increases, the labor wedge turns positive because output then rises strongly – the labor wedge becomes procyclical. See Appendix Figure 2 for the equilibrium impulse response of the labor wedge with $h = 0$, $h = 2$, and $h = 4$. The result that the labor wedge is procyclical once productivity changes may depend on the assumption of a linear disutility of labor ($\eta = 0$) which we make for computational tractability (to make it easier to find the fixed point of the model).

In Section 5 we set $\alpha = 0.33$ and $\phi = 0.65$, implying nearly constant returns to scale in capital and labor. It is interesting to ask what happens when the sum $\alpha + \phi$ is further below 1. We resolved for the RI equilibrium with $h = 2$ assuming $\alpha = 0.3$ and $\phi = 0.6$. A given mistake in the choice of capital or labor input becomes more costly (both non-zero elements in Θ, equation (8) in the paper, increase), which raises the firms' incentive to pay attention. In parallel, both profit-maximizing inputs become less responsive to productivity (the variance of each element in x_t^*, equation (9) in the paper, decreases), which lowers the firms' incentive to pay attention. We find that in equilibrium the second effect dominates, firms reduce their attention, and the impulse responses of employment and output on impact of a positive news shock are positive, like in Figure 4 in the paper.

It is interesting to compare the RI equilibrium with the PI equilibrium of the version of the model with adjustment costs. With a quadratic investment adjustment cost, firms want to smooth investment. In the PI equilibrium with $h \geq 1$, investment rises on impact of a positive news shock but consumption falls. Jaimovich and Rebelo (2009) add two further assumptions (variable capital utilization, and a new class of preferences) to obtain comovement. In the PI equilibrium with $h = 0$, the first-order autocorrelations of employment, investment, and output become positive, but the first-order autocorrelation of consumption growth becomes negative. With a quadratic labor adjustment cost, the firms' optimal labor input becomes a function of *expected future* productivity. This is a complimentary mechanism to the one in the RI equilibrium, where the optimal labor input depends on current productivity but the optimal signal confounds current and expected future productivity. We find that the PI equilibrium of the model with a quadratic labor adjustment cost is similar to the RI equilibrium analyzed thus far, for a particular choice of the parameter governing the size of the adjustment cost.[2] With $h = 0$ the first-order autocorrelations of employment, investment, and output growth become positive and are approximately in line with the data. With $h \geq 1$ the impulse responses of employment and output to a positive news shock become positive on impact, while the impulse response of investment is approximately zero. Some differences from the RI equilibrium stand out, however. The PI model with a labor adjustment cost cannot match

[2] With a quadratic labor adjustment cost, the period t profit of firm i is given by equation (2) in the paper minus a parameter times $(L_{it}/L_{it-1} - 1)^2 W_t L_{it}$. With a quadratic investment adjustment cost, on the right-hand side of the law of motion for capital (equation (1) in the paper) we subtract a parameter times $(I_{it}/K_{it-1} - \delta)^2 K_{it-1}$.

a positive coefficient in the Coibion-Gorodnichenko regression (the model predicts a coefficient of zero). Furthermore, the rational inattention model makes a specific prediction about the outcome of the comparative static experiment in Section 5.3 of the paper.

F Expected loss in utility from suboptimal actions

This appendix contains the derivation of objective (21) in the paper. Throughout this appendix, we assume that the household chooses asset holdings, q_{jt}, and hours worked, l_{jt}, in every period t.

First, using the flow budget constraint to substitute for consumption in the utility function and expressing all variables in terms of log-deviations from the non-stochastic steady state yields the following expression for the period utility of household j in period t:

$$\begin{aligned} f\left(q_{jt}, l_{jt}, q_{jt-1}, d_t, v_t, w_t\right) &= \frac{C^{1-\gamma}}{1-\gamma}\left[\omega_W e^{w_t+l_{jt}} + \omega_D e^{d_t+q_{jt-1}} - \omega_V e^{v_t}\left(e^{q_{jt}} - e^{q_{jt-1}}\right)\right]^{1-\gamma} - \frac{1}{1-\gamma} \\ &\quad - \frac{L^{1+\eta}}{1+\eta} e^{(1+\eta)l_{jt}}. \end{aligned} \tag{10}$$

The period utility of household j in period t depends on three sets of variables: variables that the household chooses in period t (q_{jt} and l_{jt}), a variable that the household chose in period $t-1$ (q_{jt-1}), and variables that the household takes as given (d_t, v_t and w_t).

Next, the following two propositions show that, after the second-order Taylor approximation of f at the non-stochastic steady state, the loss in the expected discounted sum of period utility due to suboptimal actions is given by expression (21) in the paper.

Proposition 3 *Let $E_{j,-1}$ denote the expectation operator conditioned on information of household j in period -1. Let g denote the functional that is obtained by multiplying the period utility function (10) by β^t and summing over all t from zero to infinity. Let $\tilde{g}$ denote the second-order Taylor approximation of g at the non-stochastic steady state. Let χ_t, z_t and ϱ_t denote the following vectors*

$$\chi_t = \begin{pmatrix} q_{jt} \\ l_{jt} \end{pmatrix} \quad z_t = \begin{pmatrix} d_t \\ v_t \\ w_t \end{pmatrix} \quad \varrho_t = \begin{pmatrix} \chi_t \\ z_t \\ 1 \end{pmatrix}.$$

Let $\varrho_{m,t}$ and $\varrho_{n,t}$ denote the mth element and the nth element of the vector ϱ_t.

Suppose that household j knows in period -1 its initial share in the mutual fund, $q_{j,-1}$. Suppose also that there exist two constants $\delta < (1/\beta)$ and $A \in \mathbb{R}$ such that, for each period $t \geq 0$, for all $m, n \in \{1,2,3,4,5,6\}$, and for $\tau = 0, 1$,

$$E_{j,-1} \left| \varrho_{m,t} \varrho_{n,t+\tau} \right| < \delta^t A. \tag{11}$$

Then, after the second-order Taylor approximation of f at the non-stochastic steady state, the loss in expected utility when the law of motion for the actions differs from the law of motion for the optimal actions under perfect information is given by

$$\begin{aligned} & E_{j,-1}\left[\tilde{g}\left(q_{j,-1}, \chi_0, z_0, \chi_1, z_1, \ldots\right)\right] - E_{j,-1}\left[\tilde{g}\left(q_{j,-1}, \chi_0^*, z_0, \chi_1^*, z_1, \ldots\right)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{1}{2}\left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right)\right], \end{aligned} \tag{12}$$

where the matrices Θ_0 and Θ_1 are given by

$$\Theta_0 = -C^{1-\gamma} \begin{bmatrix} \gamma\omega_V^2 \left(\frac{1}{\beta} + 1\right) & -\gamma\omega_V\omega_W \\ -\gamma\omega_V\omega_W & \omega_W\left(\omega_W\gamma + \eta\right) \end{bmatrix} \quad \Theta_1 = C^{1-\gamma} \begin{bmatrix} \gamma\omega_V^2 & -\gamma\omega_V\omega_W \\ 0 & 0 \end{bmatrix}, \tag{13}$$

and the sequence of optimal actions under perfect information, denoted $\{\chi_t^*\}_{t=0}^{\infty}$, is given by: the saving function

$$\omega_V q_{jt}^* = \omega_V q_{jt-1}^* + \zeta_t - (1-\beta)\sum_{s=t}^{\infty} \beta^{s-t} E_t\left[\zeta_s\right] + \left(1 + \omega_W \frac{\gamma}{\eta}\right) \frac{1}{\gamma} \beta \sum_{s=t}^{\infty} \beta^{s-t} E_t\left[r_{s+1}\right], \tag{14}$$

the optimality condition for labor supply

$$\gamma\left[\omega_V\left(\frac{1}{\beta} q_{jt-1}^* - q_{jt}^*\right) + \omega_W\left(w_t + l_{jt}^*\right) + \omega_D d_t\right] + \eta l_{jt}^* = w_t, \tag{15}$$

and the initial condition for the household's share in the mutual fund

$$q_{j,-1}^* = q_{j,-1}. \tag{16}$$

Here E_t denotes the expectation operator conditioned on the entire history up to and including period t, $\zeta_s \equiv \omega_W\left(1 + \frac{1}{\eta}\right) w_s + \omega_D d_s$ and $r_{s+1} \equiv \beta v_{s+1} - v_s + (1-\beta) d_{s+1}$, and ω_V, ω_W and ω_D denote the steady-state ratios V/C, WL/C and D/C, respectively.

Proof: First, the period utility of household j in period t depends on three sets of variables: variables that the household chooses in period t (q_{jt} and l_{jt}), a variable that the household chose in period $t-1$ (q_{jt-1}), and variables that the household takes as given (d_t, v_t and w_t). The first set of variables is collected in the vector χ_t, the second set of variables is an element of χ_{t-1} for all $t \geq 0$ once one defines the vector $\chi_{-1} = (q_{j,-1}, 0)'$, and the third set of variables is collected in the vector z_t. Hence, the period utility of household j in period t depends only on the vectors χ_t, χ_{t-1} and z_t.

The next six steps are word for word identical to the steps "Second" to "Seventh" in proof of Proposition 2 in online Appendix D of Maćkowiak and Wiederholt (2015). Conditions (40)-(41) in online Appendix D of Maćkowiak and Wiederholt (2015) are satisfied because of the assumption in Proposition 3 that household j knows in period -1 its initial share in the mutual fund, $q_{j,-1}$. Condition (42) in online Appendix D of Maćkowiak and Wiederholt (2015) is equal to condition (11) in Proposition 3. These steps "Second" to "Seventh" yield equation (12), where Θ_0 is defined as the Hessian matrix of second derivatives of g with respect to χ_t evaluated at the non-stochastic steady state and divided by β^t, Θ_1 is defined as the Hessian matrix of second derivatives of g with respect to χ_t and χ_{t+1} evaluated at the non-stochastic steady state and divided by β^t, and the process $\{\chi_t^*\}_{t=0}^{\infty}$ is defined as the sequence of actions that the household would take if it had perfect information in every period $t \geq 0$.

Eighth, the functional g in Proposition 3 is the discounted sum of period utility

$$g\left(\chi_{-1}, \chi_0, z_0, \chi_1, z_1, \ldots\right) = \sum_{t=0}^{\infty} \beta^t f\left(\chi_t, \chi_{t-1}, z_t\right),$$

where the function f is the period utility function

$$\begin{aligned} f\left(\chi_t, \chi_{t-1}, z_t\right) \quad = \quad & \frac{C^{1-\gamma}}{1-\gamma}\left[\omega_W e^{w_t + l_{jt}} + \omega_D e^{d_t + q_{jt-1}} - \omega_V e^{v_t}\left(e^{q_{jt}} - e^{q_{jt-1}}\right)\right]^{1-\gamma} - \frac{1}{1-\gamma} \\ & -\frac{L^{1+\eta}}{1+\eta} e^{(1+\eta) l_{jt}}. \end{aligned}$$

Computing the matrices Θ_0 and Θ_1 for this functional g yields equation (13).

Ninth, we characterize the optimal actions under perfect information. Formally, the process $\{\chi_t^*\}_{t=0}^{\infty}$ is defined by the initial condition $\chi_{-1}^* = (q_{j,-1}, 0)'$, the optimality condition

$$\forall t \geq 0 : E_t\left[\theta_0 + \Theta_{-1}\chi_{t-1}^* + \Theta_0\chi_t^* + \Theta_1\chi_{t+1}^* + \Phi_0 z_t + \Phi_1 z_{t+1}\right] = 0, \tag{17}$$

and the condition that the vector ϱ_t with $\chi_t = \chi_t^*$ satisfies condition (11). See the step "Fourth" in proof of Proposition 2 in online Appendix D of Maćkowiak and Wiederholt (2015). Here θ_0 is defined as the vector of first derivatives of g with respect to χ_t evaluated at the non-stochastic steady state and divided by β^t, Θ_{-1} is defined as the matrix of second derivatives of g with respect to χ_t and χ_{t-1} evaluated at the non-stochastic steady state and divided by β^t, Φ_0 is defined as the matrix of second derivatives of g with respect to χ_t and z_t evaluated at the non-stochastic steady state and divided by β^t, Φ_1 is defined as the matrix of second derivatives of g with respect to χ_t and z_{t+1} evaluated at the non-stochastic steady state and divided by β^t, and E_t denotes the expectation operator conditioned on the entire history up to and including period t. Computing the vector θ_0 and the matrices Θ_{-1}, Φ_0, and Φ_1 for the functional g defined in the previous step yields

$$\theta_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \Theta_{-1} = C^{1-\gamma} \begin{bmatrix} \gamma\omega_V^2 \frac{1}{\beta} & 0 \\ -\gamma\omega_V\omega_W \frac{1}{\beta} & 0 \end{bmatrix}$$

$$\Phi_0 = C^{1-\gamma} \begin{bmatrix} \gamma\omega_D\omega_V & -\omega_V & \gamma\omega_W\omega_V \\ -\gamma\omega_D\omega_W & 0 & -\gamma\omega_W^2 + \omega_W \end{bmatrix} \quad \Phi_1 = C^{1-\gamma} \begin{bmatrix} -\gamma\omega_D\omega_V + \beta\omega_D & \beta\omega_V & -\gamma\omega_W\omega_V \\ 0 & 0 & 0 \end{bmatrix}.$$

Substituting the equations for θ_0, Θ_{-1}, Φ_0, Φ_1, Θ_0, Θ_1, χ_t and z_t into equation (17) yields

$$\begin{aligned} & \omega_V \left(\frac{1}{\beta} q^*_{jt-1} - q^*_{jt} \right) + \omega_W \left(w_t + l^*_{jt} \right) + \omega_D d_t \\ = \;& E_t \left[\omega_V \left(\frac{1}{\beta} q^*_{jt} - q^*_{jt+1} \right) + \omega_W \left(w_{t+1} + l^*_{jt+1} \right) + \omega_D d_{t+1} \right] - \frac{1}{\gamma} E_t \left[r_{t+1} \right] \end{aligned} \tag{18}$$

and

$$\gamma \left[\omega_V \left(\frac{1}{\beta} q^*_{jt-1} - q^*_{jt} \right) + \omega_W \left(w_t + l^*_{jt} \right) + \omega_D d_t \right] + \eta l^*_{jt} = w_t, \tag{19}$$

where we have used the definition

$$r_{t+1} \equiv \beta v_{t+1} - v_t + (1-\beta) d_{t+1}. \tag{20}$$

Equation (18) is the combination of the usual Euler equation and the flow budget constraint. Equation (19) is the combination of the usual optimality condition for labor supply and the flow budget constraint. Equation (20) is the return on the mutual fund.

Tenth, we show that equations (18)-(19) and condition (11) imply the saving function (14). For this purpose, it is useful to write equations (18)-(19) as

$$c^*_{jt} = E_t \left[c^*_{jt+1} \right] - \frac{1}{\gamma} E_t \left[r_{t+1} \right] \tag{21}$$

and

$$\gamma c^*_{jt} + \eta l^*_{jt} = w_t, \tag{22}$$

where the variable c_{jt}^* is defined as

$$c_{jt}^* \equiv \omega_V \left(\frac{1}{\beta} q_{jt-1}^* - q_{jt}^*\right) + \omega_W \left(w_t + l_{jt}^*\right) + \omega_D d_t. \tag{23}$$

Using equation (22) to substitute for l_{jt}^* in the definition (23) and rearranging yields the equation

$$\left(1 + \omega_W \frac{\gamma}{\eta}\right) c_{jt}^* = \omega_V \frac{1}{\beta} q_{jt-1}^* - \omega_V q_{jt}^* + \omega_W \left(1 + \frac{1}{\eta}\right) w_t + \omega_D d_t, \tag{24}$$

which implies

$$\left(1 + \omega_W \frac{\gamma}{\eta}\right) \sum_{s=t}^{t+N} \beta^{s-t} c_{js}^* = \omega_V \frac{1}{\beta} q_{jt-1}^* + \sum_{s=t}^{t+N} \beta^{s-t} \left[\omega_W \left(1 + \frac{1}{\eta}\right) w_s + \omega_D d_s\right] - \omega_V \beta^N q_{jt+N}^*.$$

Taking the expectation $E_t[\cdot]$ on both sides of the last equation yields

$$\left(1 + \omega_W \frac{\gamma}{\eta}\right) \sum_{s=t}^{t+N} \beta^{s-t} E_t\left[c_{js}^*\right] = \omega_V \frac{1}{\beta} q_{jt-1}^* + \sum_{s=t}^{t+N} \beta^{s-t} E_t \left[\omega_W \left(1 + \frac{1}{\eta}\right) w_s + \omega_D d_s\right] - E_t\left[\omega_V \beta^N q_{jt+N}^*\right].$$

Taking the limit as $N \to \infty$ and using condition (11), which implies $\lim_{N\to\infty} \beta^N E_t\left[q_{jt+N}^*\right] = 0$, yields

$$\left(1 + \omega_W \frac{\gamma}{\eta}\right) \sum_{s=t}^{\infty} \beta^{s-t} E_t\left[c_{js}^*\right] = \omega_V \frac{1}{\beta} q_{jt-1}^* + \sum_{s=t}^{\infty} \beta^{s-t} E_t \left[\omega_W \left(1 + \frac{1}{\eta}\right) w_s + \omega_D d_s\right].$$

Next, using equation (21) and the law of iterated expectations yields

$$\sum_{s=t}^{\infty} \beta^{s-t} E_t\left[c_{js}^*\right] = \frac{1}{1-\beta} c_{jt}^* + \frac{1}{\gamma} \frac{1}{1-\beta} \sum_{s=t+1}^{\infty} \beta^{s-t} E_t\left[r_s\right].$$

Combining the last two equations and solving for c_{jt}^* yields

$$\begin{aligned} c_{jt}^* &= \frac{\omega_V}{1 + \omega_W \frac{\gamma}{\eta}} \frac{1-\beta}{\beta} q_{jt-1}^* + \frac{1}{1 + \omega_W \frac{\gamma}{\eta}} (1-\beta) \sum_{s=t}^{\infty} \beta^{s-t} E_t \left[\omega_W \left(1 + \frac{1}{\eta}\right) w_s + \omega_D d_s\right] \\ &\quad - \frac{1}{\gamma} \sum_{s=t+1}^{\infty} \beta^{s-t} E_t\left[r_s\right]. \end{aligned}$$

Finally, substituting equation (24) into the last equation and rearranging yields

$$\begin{aligned} \omega_V q_{jt}^* &= \omega_V q_{jt-1}^* + \left[\omega_W \left(1 + \frac{1}{\eta}\right) w_t + \omega_D d_t\right] - (1-\beta) \sum_{s=t}^{\infty} \beta^{s-t} E_t \left[\omega_W \left(1 + \frac{1}{\eta}\right) w_s + \omega_D d_s\right] \\ &\quad + \left(1 + \omega_W \frac{\gamma}{\eta}\right) \frac{1}{\gamma} \sum_{s=t+1}^{\infty} \beta^{s-t} E_t\left[r_s\right]. \end{aligned}$$

Proposition 4 *Under condition (11), we have*

$$\begin{aligned}&\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}\left(\chi_t-\chi_t^*\right)'\Theta_0\left(\chi_t-\chi_t^*\right)+\left(\chi_t-\chi_t^*\right)'\Theta_1\left(\chi_{t+1}-\chi_{t+1}^*\right)\right]\\ =\;&\sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}\left(\tilde{x}_t-\tilde{x}_t^*\right)'\tilde{\Theta}\left(\tilde{x}_t-\tilde{x}_t^*\right)\right],\end{aligned}\tag{25}$$

where $\tilde{x}_t$, $\tilde{\Theta}$ and $\tilde{x}_t^*$ are defined as

$$\tilde{x}_t=\begin{pmatrix}\omega_V\left(q_{jt}-q_{jt-1}\right)\\ \gamma\left[\omega_V\left(\frac{1}{\beta}q_{jt-1}-q_{jt}\right)+\omega_W l_{jt}\right]+\eta l_{jt}\end{pmatrix},\tag{26}$$

$$\tilde{\Theta}=-C^{1-\gamma}\gamma\begin{bmatrix}\left(1-\frac{1}{1+\frac{\eta}{\omega_W\gamma}}\right)\frac{1}{\beta} & 0\\ 0 & \frac{1}{1+\frac{\eta}{\omega_W\gamma}}\frac{1}{\gamma^2}\end{bmatrix},\tag{27}$$

and

$$\tilde{x}_t^*=\begin{pmatrix}\zeta_t-(1-\beta)\sum_{s=t}^{\infty}\beta^{s-t}E_t\left[\zeta_s\right]+\left(1+\omega_W\frac{\gamma}{\eta}\right)\frac{1}{\gamma}\beta\sum_{s=t}^{\infty}\beta^{s-t}E_t\left[r_{s+1}\right]\\ w_t-\gamma\left(\omega_W w_t+\omega_D d_t\right)\end{pmatrix}.\tag{28}$$

Proof: First, the period t term on the left-hand side of equation (25) equals

$$\frac{1}{2}\left(\chi_t-\chi_t^*\right)'\Theta_0\left(\chi_t-\chi_t^*\right)+\left(\chi_t-\chi_t^*\right)'\Theta_1\left(\chi_{t+1}-\chi_{t+1}^*\right)$$

$$=C^{1-\gamma}\begin{bmatrix}-\frac{1}{2}\gamma\omega_V^2\left(\frac{1}{\beta}+1\right)\left(q_{jt}-q_{jt}^*\right)^2\\ +\gamma\omega_V\omega_W\left(q_{jt}-q_{jt}^*\right)\left(l_{jt}-l_{jt}^*\right)\\ -\frac{1}{2}\omega_W\left(\omega_W\gamma+\eta\right)\left(l_{jt}-l_{jt}^*\right)^2\\ +\gamma\omega_V^2\left(q_{jt}-q_{jt}^*\right)\left(q_{jt+1}-q_{jt+1}^*\right)\\ -\gamma\omega_V\omega_W\left(q_{jt}-q_{jt}^*\right)\left(l_{jt+1}-l_{jt+1}^*\right)\end{bmatrix}\tag{29}$$

$$=C^{1-\gamma}\begin{bmatrix}-\frac{1}{2}\left(\gamma-\frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\right)\left(\tilde{x}_{1,t+1}-\tilde{x}_{1,t+1}^*\right)^2\\ -\frac{1}{2}\frac{\omega_W}{\gamma\omega_W+\eta}\left(\tilde{x}_{2,t}-\tilde{x}_{2,t}^*\right)^2\\ +\frac{1}{2}\left(\gamma-\frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\right)\omega_V^2\left[\left(q_{jt+1}-q_{jt+1}^*\right)^2-\frac{1}{\beta}\left(q_{jt}-q_{jt}^*\right)^2\right]\\ +\frac{1}{2}\left(\frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\frac{1}{\beta}\right)\omega_V^2\left[\left(q_{jt}-q_{jt}^*\right)^2-\frac{1}{\beta}\left(q_{jt-1}-q_{jt-1}^*\right)^2\right]\\ -\left(\frac{\gamma\omega_W}{\gamma\omega_W+\eta}\right)\omega_V\left[\left(q_{jt}-q_{jt}^*\right)\left(\tilde{x}_{2,t+1}-\tilde{x}_{2,t+1}^*\right)-\frac{1}{\beta}\left(q_{jt-1}-q_{jt-1}^*\right)\left(\tilde{x}_{2,t}-\tilde{x}_{2,t}^*\right)\right]\end{bmatrix}.\tag{30}$$

Equation (29) follows from the definition of χ_t, Θ_0 and Θ_1 in Proposition 3. Equation (30) follows from the definition of $\tilde{x}_t$ in Proposition 4. Here $\tilde{x}_{1,t}$ and $\tilde{x}_{2,t}$ denote the first element and the second element of $\tilde{x}_t$, respectively. Equation (29) is easy to verify by substituting the definition of χ_t, Θ_0 and Θ_1 into the left-hand side of equation (29). Equation (30) is easy to verify by substituting the definition of $\tilde{x}_t$ into the right-hand side of equation (30).

Second, multiplying equations (29) and (30) by β^t, summing over all $t = 0, 1, 2, \ldots, T$, and using the fact that equation (16) implies $\tilde{x}_{1,0} - \tilde{x}^*_{1,0} = \omega_V \left(q_{j0} - q^*_{j0}\right)$ yields

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \left[\frac{1}{2}\left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right)\right] \\
= \quad & C^{1-\gamma} \sum_{t=0}^{T} \beta^t \left[-\frac{1}{2}\left(\gamma - \frac{\gamma^2 \omega_W}{\gamma\omega_W + \eta}\right)\frac{1}{\beta}\left(\tilde{x}_{1,t} - \tilde{x}^*_{1,t}\right)^2 - \frac{1}{2}\frac{\omega_W}{\gamma\omega_W + \eta}\left(\tilde{x}_{2,t} - \tilde{x}^*_{2,t}\right)^2\right] \\
& + C^{1-\gamma}\beta^T \begin{bmatrix} -\frac{1}{2}\left(\gamma - \frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\right)\left(\tilde{x}_{1,T+1} - \tilde{x}^*_{1,T+1}\right)^2 \\ +\frac{1}{2}\left(\gamma - \frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\right)\omega_V^2\left(q_{jT+1} - q^*_{jT+1}\right)^2 \\ +\frac{1}{2}\left(\frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\frac{1}{\beta}\right)\omega_V^2\left(q_{jT} - q^*_{jT}\right)^2 \\ -\left(\frac{\gamma\omega_W}{\gamma\omega_W+\eta}\right)\omega_V\left(q_{jT} - q^*_{jT}\right)\left(\tilde{x}_{2,T+1} - \tilde{x}^*_{2,T+1}\right) \end{bmatrix}.
\end{aligned} \tag{31}$$

Third, taking the expectation $E_{j,-1}[\cdot]$ on both sides of the last equation and taking the limit as $T \to \infty$ yields

$$\begin{aligned}
&\sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[\frac{1}{2}\left(\chi_t - \chi_t^*\right)' \Theta_0 \left(\chi_t - \chi_t^*\right) + \left(\chi_t - \chi_t^*\right)' \Theta_1 \left(\chi_{t+1} - \chi_{t+1}^*\right)\right] \\
= \; & C^{1-\gamma} \sum_{t=0}^{\infty} \beta^t E_{j,-1} \left[-\frac{1}{2}\left(\gamma - \frac{\gamma^2 \omega_W}{\gamma\omega_W + \eta}\right)\frac{1}{\beta}\left(\tilde{x}_{1,t} - \tilde{x}^*_{1,t}\right)^2 - \frac{1}{2}\frac{\omega_W}{\gamma\omega_W + \eta}\left(\tilde{x}_{2,t} - \tilde{x}^*_{2,t}\right)^2\right] \\
& + C^{1-\gamma} \lim_{T\to\infty} \beta^T E_{j,-1} \begin{bmatrix} -\frac{1}{2}\left(\gamma - \frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\right)\left(\tilde{x}_{1,T+1} - \tilde{x}^*_{1,T+1}\right)^2 \\ +\frac{1}{2}\left(\gamma - \frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\right)\omega_V^2\left(q_{jT+1} - q^*_{jT+1}\right)^2 \\ +\frac{1}{2}\left(\frac{\gamma^2\omega_W}{\gamma\omega_W+\eta}\frac{1}{\beta}\right)\omega_V^2\left(q_{jT} - q^*_{jT}\right)^2 \\ -\left(\frac{\gamma\omega_W}{\gamma\omega_W+\eta}\right)\omega_V\left(q_{jT} - q^*_{jT}\right)\left(\tilde{x}_{2,T+1} - \tilde{x}^*_{2,T+1}\right) \end{bmatrix}.
\end{aligned} \tag{32}$$

Condition (11) implies that the second term on the right-hand side of the last equation equals zero.

Fourth, the definition (26) and equations (14)-(15) imply equation (28).

References

Afrouzi, H. and C. Yang (2021): "Dynamic Rational Inattention and the Phillips Curve," Working paper.

Jaimovich, N. and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle?" *American Economic Review*, 99, 1097–1118.

Maćkowiak, B., F. Matějka, and M. Wiederholt (2018): "Dynamic Rational Inattention: Analytical Results," *Journal of Economic Theory*, 176, 650–692.

Maćkowiak, B. and M. Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769–803.

——— (2015): "Business Cycle Dynamics under Rational Inattention," *Review of Economic Studies*, 82, 1502–1532.

Shimer, R. (2009): "Convergence in Macroeconomics: The Labor Wedge," *American Economic Journal: Macroeconomics*, 1, 280–97.

Appendix Table 1: Varying *h* in the law of motion for productivity

	Coefficient in the optimal signal on								
h	ε_t	ε_{t-1}	ε_{t-2}	ε_{t-3}	ε_{t-4}	ε_{t-5}	σ_ψ	Attention, bits per period	Labor input on impact as fraction of maximum labor input with perfect information
0	0	0	0	0	0	0	0.0191	0.216	0.26
1	0.0055	0	0	0	0	0	0.0195	0.211	0.15
2	0.0045	0.0059	0	0	0	0	0.0200	0.201	0.09
3	0.0037	0.0051	0.0063	0	0	0	0.0205	0.192	0.06
4	0.0031	0.0044	0.0056	0.0065	0	0	0.0212	0.183	0.04
5	0.0027	0.0039	0.0050	0.0060	0.0067	0	0.0219	0.175	0.03
6	0.0022	0.0034	0.0045	0.0056	0.0064	0.0069	0.0227	0.168	0.02

Productivity follows the law of motion $a_t = \rho a_{t-1} + \sigma\varepsilon_{t-h}$.

The only variable input is labor, α = 0, and parameter values are as in Section 4.1.

A measure zero of firms are subject to rational inattention. Other firms and all households have perfect information.

The table reports how the optimal signal and the labor input based on the optimal signal vary with *h* .

Appendix Table 2: Varying the persistence of productivity

ρ	Coefficient in the optimal signal on ε_t	σ_ψ	Attention, bits per period	Expected profit loss
0.5	0.0110	0.0303	0.216	2.00E-05
0.6	0.0099	0.0271	0.234	1.79E-05
0.7	0.0087	0.0244	0.242	1.55E-05
0.8	0.0074	0.0219	0.239	1.29E-05
0.9	0.0055	0.0195	0.211	9.55E-06
0.95	0.0041	0.0181	0.172	7.14E-06

Productivity follows the law of motion $a_t = \rho a_{t-1} + \sigma\varepsilon_{t-h}$ with *h* = 1.

The only variable input is labor, α = 0, and parameter values are as in Section 4.1 except as otherwise indicated.

A measure zero of firms are subject to rational inattention. Other firms and all households have perfect information.

The table reports how the optimal signal varies with ρ.

The value of *σ* is adjusted so that the unconditional variance of the optimal action is constant across the rows.

The last column gives the per period expected profit loss at the solution as a fraction of steady-state output.

Appendix Table 3: An extra term in the law of motion for productivity*

	Coefficient in the optimal signal on						
π	ε_t	ε_{t-1}	ε_{t-2}	ε_{t-3}	σ_ψ	Attention, bits per period	Labor input on impact as fraction of maximum labor input with perfect information
0	0.0031	0.0044	0.0056	0.0065	0.0212	0.183	0.04
0.002	0.0036	0.0051	0.0052	0.0062	0.0203	0.220	0.05
0.004	0.0040	0.0057	0.0046	0.0057	0.0197	0.247	0.06
0.006	0.0043	0.0062	0.0038	0.0049	0.0194	0.265	0.07
0.008	0.0046	0.0065	0.0027	0.0036	0.0192	0.270	0.08

***Productivity follows the law of motion $a_t = \rho a_{t-1} + \pi\varepsilon_{t-2} + \sigma\varepsilon_{t-4}$.**

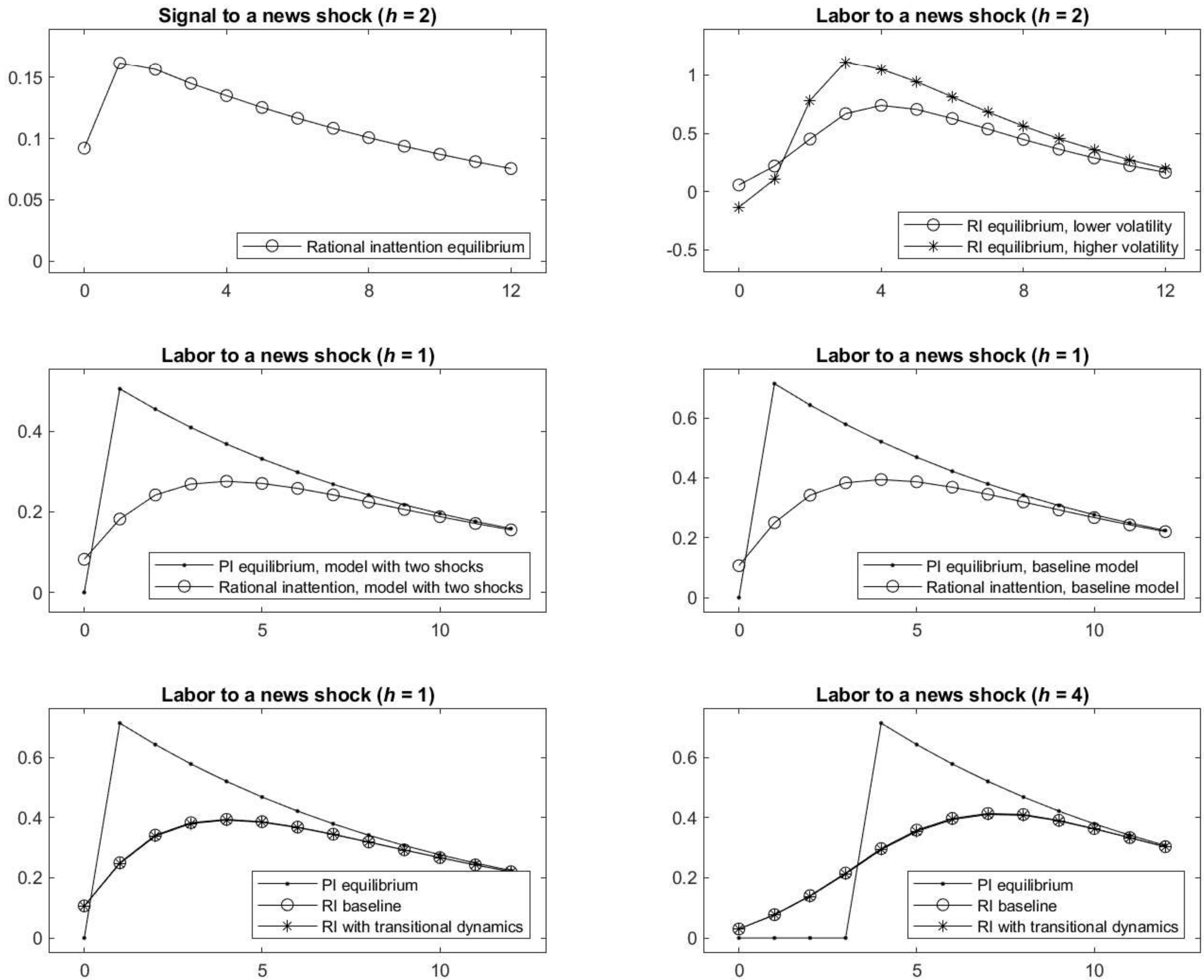
The only variable input is labor, α = 0, and parameter values are as in Section 4.1 except as otherwise indicated.

A measure zero of firms are subject to rational inattention. Other firms and all households have perfect information.

The table reports how the optimal signal and the labor input based on the optimal signal vary with *π* .

The value of *σ* is adjusted so that the unconditional variance of the optimal action is constant across the rows.

Appendix Figure 1: Additional impulse responses for Section 5 and Appendix D



Appendix Figure 2: Impulse responses for Appendix E