MACROECONOMIC SHOCKS & THEIR PROPAGATION UNDER RATIONAL INATTENTION

ETIENNE BRIAND[†]
Université du Québec à Montréal

THIS VERSION: AUGUST 2026

Abstract

We construct a medium-scale DSGE model with rational inattention. We show that to reproduce the sluggish dynamics of investment observed in the data, limited attention must be paired with a real cost of adjusting capital. Our model, estimated with quasi-Bayesian techniques, matches the macroeconomic evidence on the impulse responses to a monetary policy shock as well as a prototypical New Keynesian model, and is even favored by the marginal data density. Though untargeted, the responses to the technology and investment shocks — the main drivers of the business cycle — are alike in both models; those to the markup shocks are not. Lastly, the RI-DSGE matches firm-level evidence the NK-DSGE contradicts.

Keywords: information choice, rational inattention, business cycles, investment, prices.

JEL Codes: D83, E31, E32, E52, E71

[†] For comments and useful suggestions, we thank Alain Guay and participants at the 8th Behavioral Macroeconomics Workshop Heterogeneity and Expectations in Macroeconomics and Finance.
e-mail: etiennebriand1@gmail.com
web: etiennebriand.github.io.

1. Introduction

Medium-scale New Keynesian DSGE models *à la* Smets and Wouters (2007) and their extensions dominate the study of business cycle fluctuations.[1] Their appeal rests on computational tractability and capacity to fit the data: (i) full-information rational expectations models solve in milliseconds on a laptop, and (ii) adjustment frictions (Calvo or Rotemberg pricing, investment adjustment costs, habit formation, etc.) are flexible enough to match the persistence of prices and quantities observed in the data. However, many of these models' underlying assumptions—their microfoundations—are either implausible or in direct conflict with micro evidence. The concern is not new:

> *"My own view is that the best of these models ought to be regarded as proxies for better, harder-to-construct models . . . "* (Sims 1998, p.319).

Sims conjectured that in such a model, the sluggish responses of prices and quantities to shocks would instead arise from agents optimally allocating their limited attention. He termed the mechanism *rational inattention*.

In this paper, we formally test Sims' conjecture in a medium-scale DSGE model. Doing so requires overcoming two major difficulties: a theoretical one, building and solving what is, to the best of our knowledge, the first RI-DSGE to feature both pricing and capital decisions, and an empirical one, taking the model rigorously to the data.

We start by presenting our model, in which firms and households are both subject to rational inattention. Our framework is close to Maćkowiak and Wiederholt (2015), but is augmented with firm-owned capital and a non-linear disutility of labor, making it directly comparable to a prototypical medium-scale New Keynesian model.

The propagation of macroeconomic shocks in our model is dampened through strategic complementarities and imperfect information arising from limited attention. However, we show that in our setup these forces alone cannot account for the evidence of inertia and persistence in aggregate quantities; an additional friction affecting investment dynamics is required. We provide intuition for this result: take a contractionary demand shock that induces firms to cut prices; the slow adjustment of prices due to inattention feeds back into the demand for consumption, which amplifies optimal disinvestment. Without a real friction restraining the latter — which we implement with capital adjustment costs (Hayashi 1982) — no degree of inattention can deliver inertia in both prices and investment. This limitation of inattention had

[1] VAR and Local Projection models can address these questions in an unrestricted setting, but few structural shocks admit unambiguous identification.

until now been documented only in partial equilibrium (Zorn 2020, Miao, Wu, and Young 2022); we show that general equilibrium worsens the problem, as the feedback amplifies the investment response.

Solving the model is a challenge. In the RI-DSGE, each decision-maker chooses an optimal signal about the state of the economy, trading off expected losses from suboptimal actions against the cost of attention. The equilibrium is a fixed point: each agent's optimal signal depends on the signals chosen by all others (Maćkowiak and Wiederholt 2025). Using a guess-and-verify procedure built on the efficient solution methods of Afrouzi and Yang (2021) and Miao, Wu, and Young (2022), we solve the fixed point accurately — it is reached from any sensible starting point. Picking an appropriate one, however, cuts the time to convergence enough to make Markov chain Monte Carlo estimation feasible.

With these hurdles cleared, we investigate our model's capacity to fit macroeconomic evidence by comparing it to a New Keynesian counterpart whose empirical success is well established. To do so, we use a quasi-Bayesian procedure (Chernozhukov and Hong 2003, Inoue and Shintani 2018) that targets the impulse responses to a monetary policy shock as in Christiano, Eichenbaum, and Trabandt (2016) — moments commonly used throughout the literature to characterize the inertia and persistence of aggregate variables (Rotemberg and Woodford 1997, Christiano, Eichenbaum, and Evans 2005). We do so in two ways: (i) directly, using responses to a Romer and Romer (2004) identified monetary shock, and (ii) indirectly, using responses implied by the posterior distribution of the NK-DSGE, itself estimated on the full covariance structure of aggregate time series. Together, the two methods build a meaningful picture: the direct estimation delivers a formal test — the Bayes factor between the two models — and the indirect route extends it across the remaining business-cycle shocks.

The direct estimation finds that our model fits the empirical responses as well as the NK-DSGE, with the marginal data densities even favoring the RI-DSGE.

The indirect estimation shows that the RI-DSGE also reproduces the transmission of monetary policy implied by a structural model, and extends the comparison to the remaining shocks. Feeding the NK-estimated shock processes into the RI-DSGE yields its full set of impulse responses and covariance structure — and hence its likelihood. By construction the RI-DSGE attains a lower likelihood, but interestingly the gap loads almost entirely on the markup shocks, long criticized as reduced-form wedges (Chari, Kehoe, and McGrattan 2007). The responses to the technology and investment-efficiency shocks, which account for most of the variation in quantities, nearly coincide with the NK-DSGE's, strengthening the case that these shocks are structural (Chari, Kehoe, and McGrattan 2009). The indirect estimation also puts a number on the importance of capital adjustment costs: shutting them down at the estimated posterior, investment falls on impact by 18.8 percent of its steady-state level — thirteen times the estimated response. No plausible degree of inattention substitutes for the friction, a sharp departure from the existing rational inattention literature, in which inattention on its own suffices to generate inertial responses to monetary policy shocks (Afrouzi and Yang 2021, Song and Stern 2025).

Lastly, we study the predictions of both models in terms of firm-level responses to idiosyncratic shocks. The RI firm reproduces the micro evidence that price and investment responses are fast and peak on impact (Boivin, Giannoni, and Mihov 2009, Zorn 2020): its price adjusts by a third of the frictionless response on impact under a sectoral productivity process, and by over 90 percent under a firm-level one, while its investment response declines monotonically from impact. The mechanism is the firm's endogenous attention: at the same marginal cost, it responds promptly to volatile idiosyncratic shocks and slowly to aggregate ones. The NK firm, on the other hand, fails to reproduce the evidence because its frictions are independent of the shock's properties, so it adjusts by the same fraction whatever the disturbance. Given that these predictions rest upon the models' decision rule varying or not with the stochastic environment, we view them as evidence of the RI framework's robustness to *Lucas critique*.

All in all, our paper combines three strands of literature. First, we build on the large body of work on information frictions in macroeconomics; particularly how such deviations from perfect information can explain the real effects of monetary policy (e.g., Woodford 2003, Mankiw and Reis 2002). We do so by explicitly microfounding the information structure via rational inattention (Maćkowiak and Wiederholt 2015, 2025, Afrouzi and Yang 2021, Miao, Wu, and Young 2022) and by extending the class of RI-DSGE models to economies with capital and pricing decisions.

Second, our paper relates to the large literature on the estimation of DSGE models. Our estimation strategy for the RI-DSGE follows the limited information approach, using impulse responses as targets (Rotemberg and Woodford 1997, Christiano, Eichenbaum, and Evans 2005), but casts it as a Laplace-type estimator (Chernozhukov and Hong, 2003), which delivers quasi-posterior inference and model comparison in line with modern techniques. Our sequential approach is in the spirit of the indirect inference literature (Dridi, Guay, and Renault, 2007), albeit our strategy relies on the full-information estimation of the NK model as a first step; the design isolates the propagation mechanism from the estimated shock processes and is, we believe, of interest beyond our application.

Lastly, our paper also relates to the literature that aims to build models with propagation mechanisms consistent with both micro and macro evidence (Maćkowiak and Wiederholt 2009, Nimark 2008). Altig et al. (2011) propose that competing microfoundations be evaluated on microeconomic data; our firm-level results follow that criterion.

Layout. The remainder of the paper is structured as follows. Section 2 presents the RI-DSGE. Section 3 provides intuition for how inattention propagates shocks. Section 4 briefly describes a prototypical NK-DSGE benchmark. Section 5 lays out our estimation procedure—on the impulse responses to a monetary policy shock—and compares model fit. Section 6 extends the comparison to the full set of business cycle shocks. Section 7 discusses further key differences between the two models. Section 8 concludes.

2. RI-DSGE

In this section, we describe the RI-DSGE. We first lay out the economic environment, then formulate the attention problems, and finally define the equilibrium.

2.1. Economic Environment. Time is discrete and periods correspond to quarters. The economy is populated by a continuum $i \in [0,1]$ of monopolistically competitive firms and a continuum $j \in [0,1]$ of households. Firms own the capital stock and make investment decisions. Variables are dated to the period in which they are chosen or realized. The environment includes seven sources of exogenous fluctuations.

Households. Each household supplies a different type of labor service, consumes goods from a continuum of varieties, holds nominal bonds and trades shares in a mutual fund. Households have an infinite horizon and discount the future at rate $\beta \in (0,1)$.

The period utility function is

$$U(C_{jt}, \bar{L}_{jt}) = \exp\{\epsilon_t^c\} \left(\tfrac{C_{jt}^{1-\gamma}-1}{1-\gamma} - \varphi \tfrac{\bar{L}_{jt}^{1+\psi}}{1+\psi} \right), \tag{1}$$

where C_{jt} is consumption, a CES bundle of the continuum of varieties, $\bar{L}_{jt}$ is the household's gross labor supply, $\gamma > 0$ is the inverse of the elasticity of intertemporal substitution, $\psi \geq 0$ is the inverse of the Frisch elasticity, φ scales labor disutility, and ϵ_t^c is a risk-premium shock.

$$C_{jt} + B_{jt} + S_t Q_{jt} = \tfrac{R_{t-1}}{\Pi_t} B_{jt-1} + (D_t + S_t) Q_{jt-1} + (1+\tau_w) W_{jt} L_{jt} - T_t, \tag{2}$$

where B_{jt} denotes bond holdings, R_t the nominal interest rate, $\Pi_t := P_t/P_{t-1}$ the inflation rate, W_{jt} the wage rate, D_t profits from firm ownership, S_t the price of a mutual-fund share, Q_{jt} share holdings, T_t lump-sum taxes, τ_w a wage subsidy, and $L_{jt} = \exp\{\epsilon^z_{jt}\}\bar{L}_{jt}$ effective labor.

Household j chooses its consumption C_{jt} and its shares in the mutual fund Q_{jt}, and sets a wage rate W_{jt}, taking aggregate prices and quantities as given and subject to the labor demand schedule

$$L_{jt} = \left(\hat{W}_{jt}\right)^{-\vartheta^w_t} L_t, \qquad \vartheta^w_t = \tfrac{1+\epsilon^w_t}{\epsilon^w_t}, \tag{3}$$

where $\hat{W}_{jt} := W_{jt}/W_t$ denotes the relative wage rate, ϵ^w_t is the stochastic desired net markup driving wage markup fluctuations, and ϑ^w_t is the price elasticity of substitution.[2]

Firms. Each firm's production is a function of its labor L_{it} and capital K_{it-1}:

$$Y_{it} = \exp\{\epsilon^a_t + \epsilon^a_{it}\}K^{\phi}_{it-1}L^{\alpha}_{it}, \qquad \alpha + \phi = 1, \tag{4}$$

where ϵ^a_t and ϵ^a_{it} are aggregate and idiosyncratic productivity. The firm's capital stock evolves as

$$K_{it} = \exp\{\epsilon^i_t\}I_{it} + (1-\delta)K_{it-1}, \tag{5}$$

where I_{it} denotes investment, $\delta \in (0, 1]$ is the depreciation rate and ϵ^i_t is the aggregate investment-efficiency. Period real profits are

$$D_{it} = (1+\tau_p)\hat{P}_{it}Y_{it} - W_tL_{it} - I_{it} - \tfrac{\chi_k}{2}\left(\tfrac{K_{it}}{K_{it-1}} - 1\right)^2 K_{it-1}, \tag{6}$$

where $\hat{P}_{it} := P_{it}/P_t$ is the relative price, τ_p is a production subsidy and $\chi_k \geq 0$ governs capital adjustment costs. We allow for a non-zero capital adjustment cost (Zorn 2020, Miao, Wu, and Young 2022); Section 3 develops why it is necessary. Dividends flow into a mutual fund held by households. The mutual fund instructs each firm to value its expected sum of dividends using the discount factor $\beta^t\Lambda\big(\{C_{jt}\}_{j\in[0,1]}\big)\exp\{\epsilon^c_t\}$, where $\Lambda(\cdot)$ satisfies $\Lambda(\{C_j\}_{j\in[0,1]}) = C^{-\gamma}$ at the non-stochastic steady state. Firm i sets its price P_{it}, chooses labor L_{it}, and invests I_{it} to maximize its expected discounted

[2] Equation (3) arises when firms combine the labor types with a CES technology and choose how much of each to hire under RI. Equation (7) arises in the same way from households and firms combining the varieties into consumption and investment. Both cross-sectional decisions are omitted from the attention problems; they can be shown to be independent of the aggregate equilibrium.

dividends, taking aggregate prices and quantities as given, subject to the demand schedule for variety i:

$$Y_{it} = \left(\hat{P}_{it}\right)^{-\vartheta_t^p} Y_t, \qquad \vartheta_t^p = \tfrac{1+\epsilon_t^p}{\epsilon_t^p}, \tag{7}$$

where ϵ_t^p is the stochastic desired net markup driving price markup fluctuations, and ϑ_t^p is the price elasticity of substitution. [3]

Government. The central bank sets the nominal interest rate via a Taylor rule:

$$\tfrac{R_t}{R} = \left(\tfrac{R_{t-1}}{R}\right)^{\rho_r} \left(\tfrac{\Pi_t}{\Pi}\right)^{\phi_\pi(1-\rho_r)} \left(\tfrac{Y_t}{Y_t^*}\right)^{\phi_{y^*}(1-\rho_r)} \exp\{\epsilon_t^m\}, \tag{8}$$

where R and Π are the non-stochastic steady-state interest and inflation rate respectively, Y_t^* is output absent information frictions and markup shocks, and ϵ_t^m is a monetary policy shock.[4] The fiscal authority runs a balanced budget in real terms.

Aggregation. Aggregate quantities are integrals over the relevant continuum: $C_t = \int_0^1 C_{jt}\, dj$, $L_t = \int_0^1 L_{jt}\, dj$, $B_t = \int_0^1 B_{jt}\, dj$, $Q_t = \int_0^1 Q_{jt}\, dj$, $Y_t = \int_0^1 Y_{it}\, di$, $I_t = \int_0^1 I_{it}\, di$, $K_t = \int_0^1 K_{it}\, di$. The price and wage indexes are:

$$P_t^{1-\vartheta_t^p} = \int_0^1 P_{it}^{1-\vartheta_t^p}\, di, \qquad W_t^{1-\vartheta_t^w} = \int_0^1 W_{jt}^{1-\vartheta_t^w}\, dj. \tag{9}$$

The aggregate resource constraint is $Y_t = C_t + I_t + G\exp\{\epsilon_t^g\}$ where ϵ_t^g is government spending shock.

Shocks and Aggregation. The aggregate processes $\{\epsilon_t^a, \epsilon_t^c, \epsilon_t^i, \epsilon_t^m, \epsilon_t^g\}$ follow AR(1) processes centered at zero. The markup processes $\{\log(1+\epsilon_t^p), \log(1+\epsilon_t^w)\}$ follow ARMA(1,1) processes centered around $\{\log(1+\epsilon^p), \log(1+\epsilon^w)\}$. We abstract from idiosyncratic shocks when solving for aggregate dynamics in general equilibrium, as they sum to zero in the cross-section. We collect in the vector $\{\boldsymbol{\varepsilon_t}\}$ the realizations of aggregate shocks at time t.

[3] The production and wage subsidies correct steady-state distortions from market power: $\tau_p = \epsilon^p$ and $\tau_w = \epsilon^w$.

[4] Variables without a time subscript denote non-stochastic steady-state values throughout.

2.2. Attention Problems. We assume that attention is costly for firms and households. Formally, they optimally design signals about the state of the economy, recognizing that processing more information reduces expected losses from suboptimal actions but comes with higher attention costs. We formulate each decision-maker's problem as a dynamic Linear-Quadratic-Gaussian rational inattention (LQG-RI) problem. The targets $\boldsymbol{x}_t^*$ are the optimal actions under perfect information, derived in Appendix A; actions and targets are expressed in log-deviations from the non-stochastic steady state, denoted by lowercase letters, e.g. $x_t := \log X_t - \log X$.

General formulation. In period $t = -1$, each decision-maker ι chooses an attention strategy—a pair of matrices $\{\boldsymbol{\Gamma}, \boldsymbol{\Sigma}_{\boldsymbol{\nu}}\}$ collecting the loading and the noise of each signal—to solve:

$$\max_{\{\boldsymbol{\Gamma},\boldsymbol{\Sigma}_{\boldsymbol{\nu}}\}} \left\{ \sum_{t=0}^{\infty} \beta^t E_{\iota,-1} \left[\frac{1}{2} (\boldsymbol{x}_t - \boldsymbol{x}_t^*)' \boldsymbol{\Theta}_{\boldsymbol{\iota}} (\boldsymbol{x}_t - \boldsymbol{x}_t^*) \right] - \lambda_\iota \sum_{t=0}^{\infty} \beta^t I(\boldsymbol{\xi}_t; \boldsymbol{S}_{\iota t} | \mathcal{I}_{\iota t-1}) \right\}, \tag{10}$$

subject to the law of motion for the state $\boldsymbol{\xi}_{t+1} = \boldsymbol{F}\boldsymbol{\xi}_t + \boldsymbol{\mu}_{t+1}$ with $\boldsymbol{\mu}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}})$, the signal equation $\boldsymbol{S}_{\iota t} = \boldsymbol{\Gamma}\boldsymbol{\xi}_t + \boldsymbol{\nu}_{\iota t}$, and the optimality condition $\boldsymbol{x}_t = E[\boldsymbol{x}_t^* | \mathcal{I}_{\iota t}]$. The matrix $\boldsymbol{\Theta}_{\boldsymbol{\iota}}$ weighs the losses incurred from $\boldsymbol{x}_t - \boldsymbol{x}_t^*$, the deviation of the decision-maker's actions from their optimal values. When the problem is microfounded, as it is here, $\boldsymbol{\Theta}_{\boldsymbol{\iota}}$ is the Hessian of the decision-maker's objective with respect to its actions, evaluated at the non-stochastic steady state, and is negative definite. The scalar $\lambda_\iota > 0$ is the marginal cost of attention and $I(\boldsymbol{\xi}_t; \boldsymbol{S}_{\iota t} | \mathcal{I}_{\iota t-1})$ is the conditional mutual information between the state and the signal given past information, measuring the per-period reduction in uncertainty. Noise $\boldsymbol{\nu}_{\iota t}$, drawn from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\nu}})$, is idiosyncratic across agents and averages to zero in the cross-section, so it does not affect aggregate dynamics.[5]

We assume that paying attention to each shock is a distinct activity, so the optimal attention allocated to any given shock is independent of all others. The assumption formalizes the idea that decision-makers think structurally: attending to the current state of monetary policy is a separate activity from attending to aggregate productivity or any other source of uncertainty, each drawing on its own sources of information. It also serves tractability — under joint attention, the state vector stacks every shock and the fixed point becomes too costly to solve inside an estimation. The restriction operates across shocks, not within them: for a given shock, decision-makers still bundle information optimally across their targets. Allowing

[5] It can be shown that a signal loading on the current state plus idiosyncratic Gaussian noise is of the *optimal* form given this problem. See Maćkowiak, Matějka, and Wiederholt (2018).

decision-makers to bundle information across all shocks works with a few sources of uncertainty (Afrouzi and Yang 2021), and the confounding it induces between shocks can itself generate interesting dynamics; with many shocks, however, the same confounding may significantly distort the dynamic responses, as decision-makers attribute most fluctuations to the dominant source. Quantifying the distortion — if the joint model can be solved at all — is left to future research.

Firms' losses. The firm's vector of actions is $\boldsymbol{x}_{it} = (\zeta_{it}, \Delta\tilde{k}_{it})'$, where ζ_{it} is its price corrected for its inherited capital stock and $\Delta\tilde{k}_{it}$ a weighted difference between its current and lagged capital stock. The Hessian is:

$$\boldsymbol{\Theta}_i = -C^{-\gamma} Y \begin{bmatrix} \vartheta^p\left(1 + \frac{1-\alpha}{\alpha}\vartheta^p\right) & 0 \\ 0 & \tilde{\Theta}_0^k \end{bmatrix}, \tag{11}$$

where $\tilde{\Theta}_0^k$ is increasing in the capital adjustment cost χ_k. The Hessian $\boldsymbol{\Theta}_i$ is diagonal, and its two terms govern the relative stakes of the pricing and capital decisions. For the derivation and the expressions of $\boldsymbol{x}_{it}^*$, see Appendix B.

Households' losses. The household's vector of actions is $\boldsymbol{x}_{jt} = (c_{jt}, w_{jt})'$.[6] The Hessian is:

$$\boldsymbol{\Theta}_j = -C^{1-\gamma} \begin{bmatrix} \gamma & 0 \\ 0 & \vartheta^w \omega_W (1 + \vartheta^w \psi) \end{bmatrix}, \tag{12}$$

where $\omega_W := WL/C$ is the steady-state labor income share. For the derivation and the expressions of $\boldsymbol{x}_{jt}^*$, see Appendix C.

2.3. Equilibrium. We define the equilibrium of the economy in a neighborhood of its non-stochastic steady state, consistent with the formulation of the attention problems. Given stochastic sequences for the aggregate shocks $\{\boldsymbol{\varepsilon}_t\}$ and initial conditions — $k_{i,-1}$ for every firm $i \in [0,1]$, $b_{j,-1}, q_{j,-1}$ for every household $j \in [0,1]$, and initial beliefs for every decision-maker — an equilibrium consists of (i) an attention strategy for each firm and each household, (ii) stochastic sequences of firm actions — prices $\{p_{it}\}$ and capital $\{k_{it}\}$ — and of household actions — consumption $\{c_{jt}\}$, wages $\{w_{jt}\}$, bond holdings $\{b_{jt}\}$ and share holdings $\{q_{jt}\}$ — and (iii) stochastic sequences of aggregate prices and quantities $\{p_t, w_t, \pi_t, s_t, r_t, y_t, l_t, c_t, k_t, i_t\}$, such that: (a) each firm's actions are optimal given its information set $\mathcal{I}_{it}$, and its attention strategy solves Equation (10); (b) each household's actions are optimal given its information

[6] Households set the real wage. Letting them set the nominal wage instead would make the price level an additional source of uncertainty.

set $\mathcal{I}_{jt}$, and its attention strategy solves Equation (10); (c) the monetary authority follows the Taylor rule Equation (8) and the fiscal authority runs a balanced budget in real terms; (d) goods, labor, bond and share markets clear; (e) individual prices and wages aggregate to the price and wage indexes; and (f) the perceived and actual laws of motion coincide.[7]

Appendix D defines the initial conditions, and describes the iterative procedure we use to compute the equilibrium, including the algorithm that solves the attention problems and recovers the law of motion of the state.

3. Endogenous Attention and the Propagation of Shocks

This section provides intuition for how inattention propagates shocks. We first describe what decision-makers choose to track and how their choices shape their beliefs about the economy. We then decompose the effects of inattention in general equilibrium. Lastly, we examine how inattention and capital adjustment costs interact in the transmission of monetary policy shocks — the impulse responses on which the model is estimated — and explain why capital adjustment costs are an essential feature of the model.

3.1. Attention and Beliefs. Attention is a signal-design problem: decision-makers choose what to observe and how precisely; an additional unit of information is acquired only insofar as its benefit exceeds its cost. The benefit of reducing the expected deviation of an action $\boldsymbol{x}_{\iota t}$ from its target $\boldsymbol{x}^*_{\iota t}$ is the product of two things: what a mistake costs, and how unpredictable the target is. The first is the curvature of the loss, the corresponding entry of $\boldsymbol{\Theta}_\iota$ in Equation (10); the second is the conditional variance of $\boldsymbol{x}^*_\iota$ determined by its stochastic process. Neither alone drives attention. A decision whose mistakes are costly may be tracked loosely if its target barely moves; conversely, one whose mistakes are cheap may be tracked closely if its target is volatile enough.

That intuition would be exact if decision-makers were restricted to signals of the form target-plus-noise. However, a signal can be any linear combination of the state, and bundling information across correlated targets buys more than tracking each in isolation. What is priced is therefore a direction in the state space, not an individual target. Afrouzi and Yang (2021) make this precise. They show that relevant directions are the orthogonal dimensions of the matrix $\boldsymbol{\Sigma}^{1/2}_{t|t-1}\boldsymbol{\Omega}_{\iota t}\boldsymbol{\Sigma}^{1/2}_{t|t-1}$, where the benefit matrix $\boldsymbol{\Omega}_{\iota t}$ augments the period Hessian $\boldsymbol{\Theta}_\iota$ with the discounted continuation

[7] That is, the state dynamics $\{\boldsymbol{F}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}}\}$ against which firms and households design their signals are the ones their signal choices generate.

value of information. Knowledge about a persistent state remains useful tomorrow, so information about a persistent shock is worth more than about a transitory one of the same variance. The eigenvalues of that matrix are the marginal values of information in each dimension — the same product as above, with stakes carried by $\mathbf{\Omega}_{\iota t}$ and unpredictability by $\mathbf{\Sigma}_{t|t-1}$ — and its eigenvectors are the directions along which a signal of given precision buys the most. The allocation is then a reverse water-filling problem against the marginal cost λ_ι. On the extensive margin, dimensions whose eigenvalue falls short of λ_ι receive no signal at all and the decision-maker's posterior in that direction is simply its prior; on the intensive margin, a dimension whose eigenvalue exceeds λ_ι is tracked with a signal whose precision grows with the magnitude of that eigenvalue. The number of signals a decision-maker runs is at most its number of actions, so at most two here.

Thus, our diagonalization of the Hessians, Equation (11) and Equation (12), separates the losses, but a single attention problem distorts/filters both actions away from their optimal paths and restricts what dynamics can be generated — unlike in the New Keynesian model, where separate parameters govern the sluggishness of each action. Beliefs about the targets—and about the state of the economy in general—are a function of the optimal signals. A target with little at stake is therefore still tracked closely when it comoves with one that matters; e.g., households may care only about their wage-setting decision yet still learn about optimal consumption because the two are highly correlated. The same holds for any aggregate variable; e.g., a firm ends up informed about the aggregate wage when it covaries with the directions that matter for its optimal price and capital, and uninformed about aggregates that do not.

3.2. Decomposing the Effects of Inattention. Inattention shapes equilibrium dynamics through two channels. The first is *direct*: for a given stochastic process governing the targets, a higher marginal cost of attention λ_ι implies noisier signals, less weight on new information, and actions that track their targets more sluggishly. The direct channel is the partial-equilibrium intuition, and on its own it says that more inattention means more inertia.

The second is *indirect* and runs through the endogeneity of the targets, which are themselves general-equilibrium objects. Sluggish actions change aggregate prices and quantities, and those feed back into the laws of motion for optimal actions. A target can therefore become more volatile precisely because it is being tracked less well by its decision-maker. Since attention responds to the volatility of the targets, that amplification draws more attention, partly or even fully undoing the direct effect of inattention.

3.3. Monetary Policy Shocks under Rational Inattention. We now focus on the implications of how decision-makers allocate attention for general equilibrium dynamics in the case of monetary policy shocks. In particular, we show that, contrary to previous results in the literature (see Maćkowiak and Wiederholt 2015, Afrouzi and Yang 2021, Song and Stern 2025), rational inattention on its own cannot account for the sluggish responses of prices and quantities. The reason is that, in a model with capital, slow prices imply a larger contraction of the economy, which in turn makes the optimal investment more volatile — and a volatile target draws attention.

We start with the optimality condition for prices which, unlike that of capital, takes the same form whether or not the capital adjustment cost is present. Dropping all other shocks, a firm's optimal price under perfect information is

$$p^*_{it} = p_t + \left(1 + \frac{(1-\alpha)\vartheta^p}{\alpha}\right)^{-1} \left[w_t + \frac{1-\alpha}{\alpha} y_t - \frac{\phi}{\alpha} k^*_{it-1}\right], \tag{13}$$

which the pricing target ζ^*_{it} tracks up to the inherited-capital term. The form of the second target depends on whether the capital adjustment cost is positive: the capital stock k^*_{it} when $\chi_k = 0$, the composite $\Delta\tilde{k}^*_{it}$ when $\chi_k > 0$.

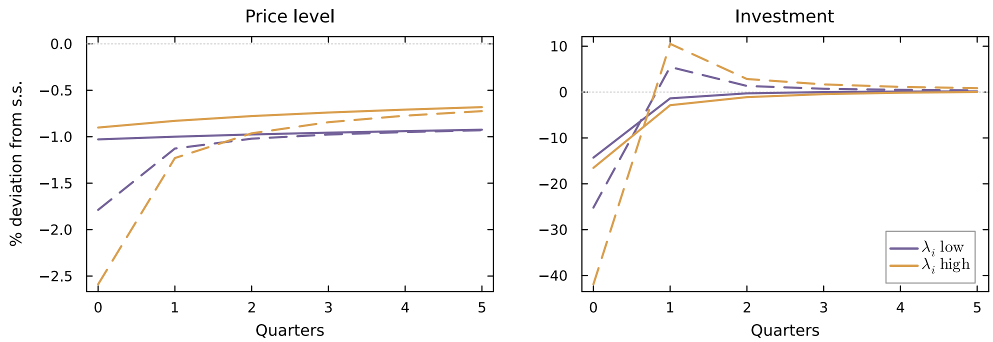
The case without capital adjustment costs. To begin, we consider the economic environment wherein inattention is the only source of inertia. With $\chi_k = 0$, the optimal capital stock under perfect information is

$$k^*_{it} = E_{it}\left\{\left(\frac{\alpha}{\alpha+\phi}\right)\left[\frac{\gamma(c_t - c_{t+1})}{1-\beta(1-\delta)} + w_{t+1} + \frac{1}{\alpha} y_{t+1}\right]\right\}. \tag{14}$$

Evaluating the structural parameters at their calibrated values ($\alpha = 2/3$, $\phi = 1/3$, $\beta = 0.99$, $\vartheta^p = 6.0$), the terms in Equation (11) equal 24.0 and 0.12: pricing mistakes are two orders of magnitude more costly than capital mistakes. All else equal, one might conjecture that firms would then devote most of their attention to tracking their optimal price, leaving the capital target loosely tracked — and investment sluggish for free.

Figure 1 depicts why that conjecture fails. Trace the contractionary monetary policy shock through the economy. The nominal rate rises and optimal prices fall. Because firms are inattentive, the price level falls only sluggishly, so the real rate rises; consumption and output contract; and the return to capital, the product of marginal cost and the marginal product of capital, falls with them. The optimal capital stock drops sharply — far more than optimal prices do. The asymmetry is written in the coefficients of the optimality conditions: strategic complementarity scales the response of p^*_{it} to aggregate conditions by $\left(1 + \frac{(1-\alpha)\vartheta^p}{\alpha}\right)^{-1} = 0.25$, while k^*_{it} loads on future output one for one and on consumption growth with a weight of

Figure 1: Firms' responses to a monetary policy shock without capital adj. cost



Note: Solid lines are the rational-inattention actions; dashed lines the perfect-information optimal actions evaluated at each equilibrium's fixed point. No capital adjustment cost ($\chi_k = 0$); firms' marginal cost of attention $\lambda_i = 120$ (low) and $\lambda_i = 400$ (high); households' $\lambda_j = 120$. The monetary policy shock follows the NK posterior-mean process; all remaining parameters are at their calibrated values.

order $\gamma/(1-\beta(1-\delta)) \approx 28.78$. Attention is allocated to the product of stakes and volatility, and what capital lacks in the first it more than makes up in the second, so firms track k^*_{it} closely and investment moves almost one for one with its target.

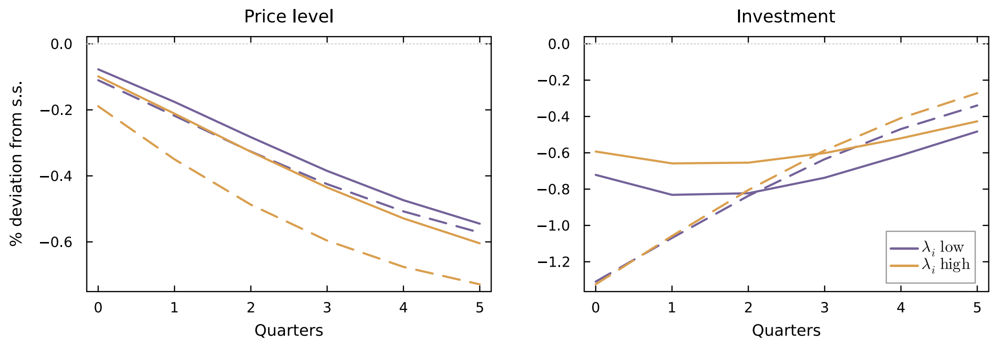
The result is a tug-of-war that inattention cannot win. Making firms more inattentive slows the price level further, which amplifies k^*_{it} further, which raises the attention devoted to capital. As the marginal cost of attention grows, investment turns more volatile and monotonic rather than hump-shaped: no degree of inattention generates inertial inflation and hump-shaped investment at the same time.

Moreover, notice that in Figure 1 firms are not even sufficiently inattentive to generate persistent inflation dynamics; that would require raising λ_i even more and making investment even more volatile. Previous RI-DSGE models had not encountered the conundrum because capital was absent.[8] Zorn (2020), Maćkowiak, Matějka, and Wiederholt (2018), and Miao, Wu, and Young (2022) solve partial equilibrium models with capital and hinted at the necessity of capital adjustment costs: loose tracking of the capital target is generally insufficient for investment to take on a hump-shaped response. In general equilibrium, the question was open until now — and the answer is that the feedback runs the other way: far from slowing investment, equilibrium forces amplify it.

The case with capital adjustment costs. We now show that including capital adjustment costs resolves the tension highlighted above; with $\chi_k > 0$, the optimal

[8] Maćkowiak and Wiederholt (2025) solve an RBC model with RI; however, firms are in perfect competition and face only TFP shocks.

Figure 2: Firms' responses to a monetary policy shock with capital adj. cost



Note: Solid lines are the rational-inattention actions; dashed lines the perfect-information optimal actions evaluated at each equilibrium's fixed point. Capital adjustment cost $\chi_k = 45$; firms' marginal cost of attention $\lambda_i = 70$ (low) and $\lambda_i = 300$ (high); households' $\lambda_j = 120$. The monetary policy shock follows the NK posterior-mean process; all remaining parameters are at their calibrated values.

capital stock solves

$$\chi_k \left[(k^*_{it} - k^*_{it-1}) - \beta E_{it}(k^*_{it+1} - k^*_{it}) \right] = E_{it}\Big[\gamma(c_t - c_{t+1}) + \big(1 - \beta(1-\delta)\big)\Big(w_{t+1} + \tfrac{1}{\alpha} y_{t+1} - \tfrac{\alpha+\phi}{\alpha}\, k^*_{it} \Big) \Big] \tag{15}$$

rather than Equation (14). The adjustment cost moves the two determinants of attention in opposite directions: it raises the stakes of the capital decision and dampens the volatility of its target.

Figure 2 shows how increasing inattention now slows down the dynamics of investment. Both determinants of attention are at work. The adjustment cost raises $\tilde{\Theta}^k_0$ because a capital mistake can no longer be undone for free: unwinding it next period is itself costly, so a given deviation from the target hurts more. The effect is quantitatively large — χ_k enters divided by $1 - \beta(1-\delta)$, which is small at quarterly frequency — and the two entries of $\mathbf{\Theta}_i$ change order above $\chi_k \approx 2.7$: capital mistakes then cost more than pricing mistakes, reversing the ranking of Equation (11). But the adjustment cost also stills the target. Equation (15) is a two-sided smoother: innovations to fundamentals are spread over many periods of Δk^*_{it}, and the surprise component of the target — the only part attention is for — shrinks. Adjustment costs thus convert the firm's capital problem from a volatile target with negligible stakes into a quiet target with substantial ones, and attention to capital need not be monotone in χ_k.

Notice that the price dynamics are now also generally more sluggish, which stems from the firm's capital stock and aggregate demand being less volatile. More inattentive firms may nonetheless induce a larger reaction in prices, as in Figure 2: the effect is entirely a general equilibrium one, operating through the demand side —

a higher λ_i implies a deeper contraction in aggregate demand, and hence a lower optimal price.

The friction in the capital accumulation breaks the tug-of-war between the pricing and capital decisions: the adjustment cost makes k_{it}^* slow, and inattention then keeps k_{it} from tracking even that slower path, so investment responds with a hump. Neither friction delivers it alone: adjustment costs alone produce monotone partial adjustment under perfect information, and inattention alone produces an amplified monotonic response. The adjustment cost helps the pricing side as well, by making the inherited-capital term in p_{it}^* slow and forecastable, so that the volatile part of the price target is dominated by aggregate news. Section 5.5 provides the quantitative counterpart of the argument.

Hence, rational inattention as a unifying explanation for the sluggishness observed in aggregate dynamic responses has its limits. However, a single friction on capital accumulation suffices to recover the qualitative results of the literature: inattention then replicates the inertia of prices and quantities from one primitive, the marginal cost of attention. Moreover, a cost of adjusting capital is a friction widely viewed as structural, accepted even among neo-classical economists: the capital stock is generally regarded as slow-moving, as in the time-to-build literature (Kydland and Prescott (1982)).

4. NK-DSGE

In this section, we define a prototypical NK-DSGE that we compare to the RI-DSGE. The two models differ in the following aspects. First, firms and households have perfect information. Second, the NK-DSGE features four adjustment frictions: prices and wages are sticky due to nominal rigidities, and consumption and investment are inertial due to habit formation and investment adjustment costs, whereas the RI-DSGE features a single one — the cost of adjusting capital.[9]

We work directly with the linearized optimality conditions. Consumption is governed by the Euler equation and by the marginal utility of consumption under habit formation,

$$\Lambda_t = E_t[\Lambda_{t+1}] - \tfrac{1}{\gamma} E_t[r_t - \pi_{t+1}] \tag{16}$$

$$\Lambda_t = \tfrac{\gamma}{(1-h)(1-\beta h)} \Big[h c_{t-1} - (1+\beta h^2) c_t + \beta h E_t[c_{t+1}] + (1-\beta h \rho_c)\epsilon_t^c \Big], \tag{17}$$

[9] The RI-DSGE cannot accommodate investment adjustment costs. The firm would track Δi_t^*, leaving capital non-stationary under rational inattention. Conversely, capital adjustment costs do not produce hump-shaped investment responses in the NK model.

the capital stock by

$$k_t = \tfrac{1}{1+\beta}\big(k_{t-1} + \beta E_t[k_{t+1}]\big) + \tfrac{1-\beta(1-\delta)}{\chi_i(1+\beta)} E_t\big[w_{t+1} + mpk_{t+1} - mpl_{t+1}\big] \\ - \tfrac{1-\beta(1-\delta)\rho_i}{\chi_i(1+\beta)}\epsilon^i_t + \tfrac{1}{\chi_i(1+\beta)}\big(E_t[\Lambda_{t+1}] - \Lambda_t\big), \quad (18)$$

and prices and wages by the two Phillips curves

$$\pi_t = \kappa_p\, mc_t + \beta E_t[\pi_{t+1}] + \hat{\epsilon}^p_t \quad (19)$$

$$w_t = \kappa_w\, \widehat{mrs}_t + \tfrac{1}{1+\beta}\big(w_{t-1} - \pi_t + \beta E_t[w_{t+1}] + \beta E_t[\pi_{t+1}]\big) + \hat{\epsilon}^w_t, \quad (20)$$

where h is the habit parameter, χ_i the investment adjustment cost, and κ_p and κ_w the slopes of the price and wage Phillips curves. The complete list of equilibrium conditions — including the definitions of marginal cost, the marginal rate of substitution gap and the marginal products — is given in Appendix E.

Writing the model at this level of generality is deliberate. At the first order, Equation (19) and Equation (20) take the same form under Calvo or Rotemberg pricing; only the mapping from κ_p and κ_w to the underlying structural parameters changes. The same is true whether firms own their capital, as in the RI-DSGE, or rent it from households, as assumed in most of the literature.[10] The same holds for Equation (18), which is unaffected by whether the adjustment cost is embedded in the capital accumulation equation *à la* Christiano, Eichenbaum, and Evans (2005) or levied in the budget constraint. We therefore estimate κ_p, κ_w, h and χ_i directly, and none of our results depends on the particular microfoundation behind them.

5. Estimation

In this section we estimate the RI- and NK-DSGE. To begin, we calibrate the structural parameters common to both models and lay out the quasi-Bayesian criterion — the distance between a model's impulse responses to a monetary policy shock and an empirical counterpart. Next, we estimate both models directly on impulse responses recovered from local projections. Lastly, we indirectly estimate the RI-DSGE on impulse responses implied by the posterior of the NK-DSGE, itself estimated on the full covariance structure of aggregate time series. The data used throughout and their construction are detailed in Appendix F.

[10] When firms own their capital, Calvo pricing depends on infinite future sequences (Sveen and Weinke 2003), but Altig et al. (2011) show that the standard NKPC still obtains, with κ_p mapping differently into the underlying price-setting frictions.

Table 1: Calibrated and estimated parameters

Panel A: Calibrated Parameters Parameter			*Panel B: Estimated Parameters* Parameter		Direct	Indirect
β	Discount factor	0.99	*B1. RI-DSGE*			
γ	EIS	1.	λ_i	Firms' marginal cost of attention	80	188
ψ	Inverse Frisch	1.	λ_j	Households' marginal cost of attention	104	218
α	Labor share	2/3	χ_k	Capital adjustment cost	56.2	18.5
ϕ	Capital share	1/3	*B2. NK-DSGE*			
δ	Depreciation of capital	0.025	κ_w	Wage NKPC	0.004	0.007
$(1+\epsilon^p)$	Price markup	1.2	κ_p	Price NKPC	0.031	0.054
$(1+\epsilon^w)$	Wage markup	1.2	χ_i	Investment adjustment cost	7.03	3.96
G/Y	Govt spending ratio	0.2	h	Habit formation	0.81	0.87
ρ_r	Taylor rule inertia	0.9	*B3: Monetary Policy Shock*			
ϕ_π	Taylor rule inflation coeff.	1.5	$100\sigma_m$	S.D. of monetary policy shock	0.163	0.220
ϕ_{y^*}	Taylor rule output gap coeff.	0.125	ρ_m	Persistence of monetary policy shock	0.0	0.044

Note: "Direct" reports the quasi-posterior medians obtained on the empirical impulse responses of Section 5.3, "Indirect" those obtained on the NK-DSGE posterior responses of Section 5.4; the NK-DSGE and monetary-policy columns under "Indirect" are the full-information posterior medians of Appendix I. The direct RI-DSGE estimates are those under the IRF-matched prior, the specification favored by the marginal data density (Table 2). In the direct estimation the monetary policy process is the one implied by the Romer–Romer shock and is held fixed; the indirect procedure re-estimates (ρ_m, σ_m) jointly with the attention parameters under the NK posterior as prior, and they barely move (ρ_m: 0.048; $100\sigma_m$: 0.214).

5.1. Calibration. We set the discount factor $\beta = 0.99$, the inverse of the elasticity of intertemporal substitution $\gamma = 1$, the inverse Frisch elasticity $\psi = 1$, the labor share $\alpha = 2/3$, the capital share $\phi = 1/3$, and the depreciation rate $\delta = 0.025$. The steady-state gross markup is 1.2 for both prices and wages. The Taylor rule parameters are the interest rate smoothing coefficient $\rho_R = 0.9$, the inflation response $\phi_\pi = 1.5$, and the output gap response $\phi_{y^*} = 0.125$, and the government spending ratio is $G/Y = 0.2$.

5.2. Quasi-Bayesian Impulse Response Matching. Let $\boldsymbol{\theta}$ denote the parameters left free in each model; i.e., the attention costs λ_i, λ_j and the capital adjustment cost χ_k, in the RI-DSGE, and the habit persistence h, the slopes of the Phillips curves κ_p, κ_w and the investment adjustment cost χ_i in the NK-DSGE. Let $\widehat{\boldsymbol{\mathcal{J}}}$ be a vector of impulse responses to a monetary policy shock and $\Sigma_{\widehat{\boldsymbol{\mathcal{J}}}}$ its covariance matrix. Following the literature, we take $\widehat{\boldsymbol{\mathcal{J}}}$ as our estimation target: our estimator $\widehat{\boldsymbol{\theta}}$ solves the minimum-distance problem[11]

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \Psi(\boldsymbol{\theta}), \qquad \Psi(\boldsymbol{\theta}) := \big(\boldsymbol{\mathcal{J}}(\boldsymbol{\theta}) - \widehat{\boldsymbol{\mathcal{J}}}\big)' \Sigma_{\widehat{\boldsymbol{\mathcal{J}}}}^{-1} \big(\boldsymbol{\mathcal{J}}(\boldsymbol{\theta}) - \widehat{\boldsymbol{\mathcal{J}}}\big). \tag{21}$$

Rather than optimize Equation (21) directly, we take a quasi-Bayesian approach: the transformation $\exp\{-\Psi(\boldsymbol{\theta})/2\}$ takes the place of a likelihood, and combined with

[11] See Rotemberg and Woodford (1997), Christiano, Eichenbaum, and Evans (2005).

a prior $\pi(\boldsymbol{\theta})$ it defines the quasi-posterior $\propto \exp\{-\Psi(\boldsymbol{\theta})/2\}\,\pi(\boldsymbol{\theta})$, a proper density over the parameter space, which we explore by Metropolis–Hastings. Chernozhukov and Hong (2003) provide the econometric foundation for this class of Laplace-type estimators; Christiano, Trabandt, and Walentin (2010) and Christiano, Eichenbaum, and Trabandt (2016) popularized its use for impulse response matching in DSGE models. The weighting matrix determines how the quasi-posterior is to be read. With $\Sigma_{\widehat{\mathcal{J}}}$ the full covariance of the targets, $\exp\{-\Psi/2\}$ is their Gaussian quasi-likelihood, and the quasi-posterior maps the uncertainty in $\widehat{\mathcal{J}}$ into uncertainty about $\boldsymbol{\theta}$.

Nothing in Equation (21) restricts where $\widehat{\mathcal{J}}$ comes from. We exploit this below by running the same procedure against two different targets. Under both, $\widehat{\mathcal{J}}$ collects the responses of the same four variables — consumption, investment, the real wage and the price level — which, as Appendix D shows, are a minimal representation of the equilibrium in the sense that the responses of every other aggregate follow from them and the exogenous processes.

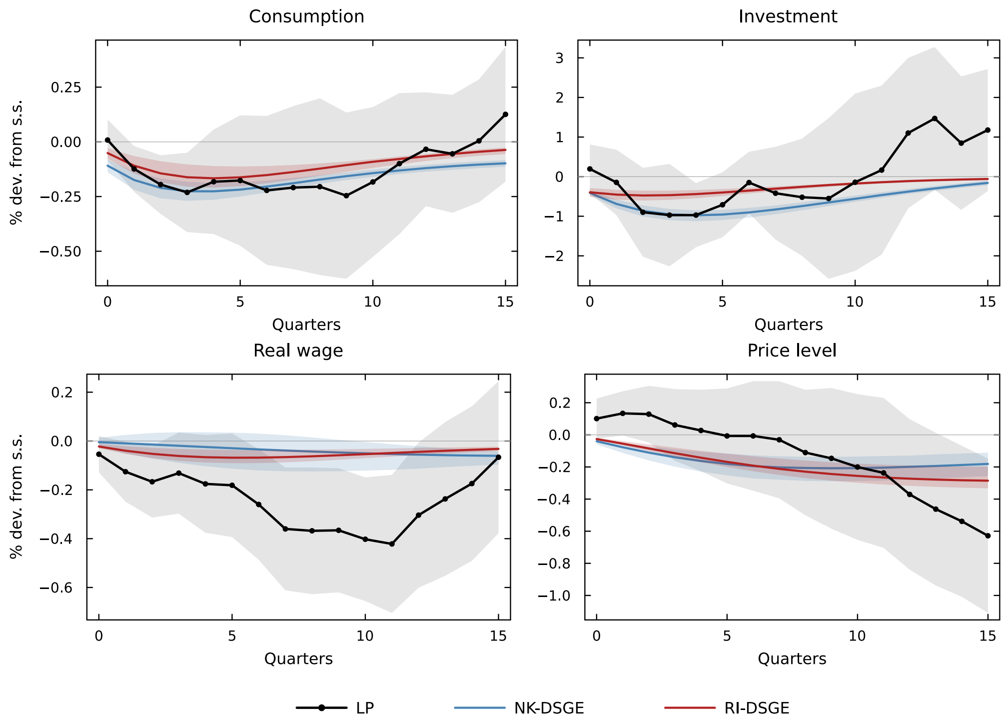
Every quasi-Bayesian estimation is run the same way. We first locate the mode of the quasi-posterior and compute the finite-difference Hessian of its logarithm there, which shapes the initial proposal. A single adaptive chain of 10,000 draws — the first 4,000 discarded as burn-in, the proposal continuously reshaped with the chain's running empirical covariance and rescaled toward an acceptance rate of 0.30 — refines it. Eight independent chains of 9,500 draws — the first 1,500 discarded — started from over-dispersed points and using the adaptive chain's empirical covariance as their proposal, deliver the 64,000 pooled draws behind the quasi-posteriors we report. The computational burden is solving the model: every draw requires computing the fixed point at the proposed parameters, warm-started from one fixed reference solution.[12] Nothing beyond that demands specialized hardware — each chain runs on a single core, and the final chains can be run in parallel — so the estimation can in principle be run on a laptop. We run all chains on a compute cluster: a solve takes between 30 and 90 seconds depending on the variant, which puts a single chain at three to seven days.[13]

5.3. Direct Estimation. Our first set of targets are impulse responses to the narrative monetary policy shock of Romer and Romer (2004) estimated by local projection (Jordà, 2005), on their original sample, 1969m3–1996m12. Aggregated to the quarterly frequency and scaled to a one-standard-deviation shock, the responses

[12] The reference solution is the mode's. Warm-starting from the last accepted draw makes the criterion path-dependent at any finite solver tolerance, which invalidates the acceptance ratio; a tight enough tolerance would resolve this — the fixed point appears unique where it exists — but at a prohibitive cost per draw.

[13] Timings are per core on a supercomputer cluster (AMD EPYC Rome, 2.2 GHz).

Figure 3: Empirical impulse responses to a monetary policy shock vs. model fit



Note: Black dotted lines: local-projection responses to a one-standard-deviation Romer and Romer (2004) shock, with 90% confidence bands in grey. Colored lines and bands: quasi-posterior median responses and 90% credible bands of the RI-DSGE (red) and NK-DSGE (blue), under the IRF-matched prior.

and their sampling covariance deliver $\widehat{\mathcal{J}}$ and $\Sigma_{\widehat{\mathcal{J}}}$; Appendix G provides the details. Because $\Sigma_{\widehat{\mathcal{J}}}$ is the full sampling covariance of the targets, the quasi-posterior admits the Gaussian reading above, and the RI- and NK-DSGE can be compared by their marginal data densities.

Marginal data densities are prior sensitive: a model that is a priori better fitting than its rival would be favored for reasons that have nothing to do with the data. The concern is real here, since the NK-DSGE receives the standard priors of the literature while the RI parameters have no such convention. We therefore discipline the RI priors by the implications of the NK ones — matching either the prior-predictive impulse responses or the prior-predictive distribution of $\Psi(\boldsymbol{\theta})$ — and report marginal data densities under both. Appendix H details the construction.

Figure 3 displays the empirical responses together with the RI and NK models' quasi-posterior responses. To the eye the fit is comparably good: both models remain inside the sampling bands throughout, the NK-DSGE digging a somewhat deeper

Table 2: Marginal Data Densities and Bayes Factors

	log-MDD	log BF
NK-DSGE	50.75	–
RI-DSGE (IRF-matched prior)	78.62	27.87
RI-DSGE (Ψ-matched prior)	62.96	12.21

Notes: Geweke modified harmonic-mean estimates with truncation $\tau = 0.5$; varying τ between 0.1 and 0.99 moves the log-MDDs by less than one log point. Log Bayes factor relative to the NK-DSGE.

investment trough and flattening its price decline where the data keep falling.

The picture is unchanged under the Ψ-matched prior: Figure 14 is nearly indistinguishable from Figure 3, both RI models fitting the same features and missing the same real-wage decline. however, are not identical (Figures 15 and 16). The posteriors depart sharply from the priors in both cases: the posterior of λ_j concentrates far above its prior — at 104 against a prior centered at 37 under the IRF-matched construction, and at 84 under the Ψ-matched one despite its deliberately tight prior — and χ_k ends up above both prior centers. The two constructions settle on different parameter combinations — $(\lambda_i, \lambda_j, \chi_k) = (80, 104, 56)$ against $(53, 84, 65)$ — but deliver similar impulse responses, the larger adjustment cost compensating for the smaller marginal costs of attention. The data are thus informative about which combinations fit — that is what moves the posteriors off the priors — but less so about the division of the fit among the three parameters, where the prior retains its influence. The fit and the Bayes factors are robust across the two priors.

Marginal data densities. Table 2 reports the result: the data favor the RI-DSGE decisively, by 27.9 log points under the IRF-matched prior and 12.2 under the conservative Ψ-matched one. Given how similar the fitted responses look, the margin comes only partly from fit. A Laplace approximation decomposes the log-MDD into the quasi-likelihood at the posterior mode, the prior density at the mode, and the volume of parameter space over which the fit survives; the comparison loads on the last two. The RI-DSGE matches the targets at parameter values its prior considers ordinary and over a wide region of its parameter space, whereas the NK-DSGE needs several parameters in the tails of its own prior ($h = 0.81$, $\kappa_w = 0.004$, $\chi_i = 7.0$) — a prior-plausibility penalty of roughly -10.6 log points on its own — and holds its fit only in that narrow configuration.

5.4. Indirect Estimation. Our second set of targets are the impulse responses to a monetary policy shock implied by the posterior of the NK-DSGE, itself estimated on the full covariance structure of aggregate time series. The exercise doubles as a robustness check on the direct estimation: the monetary policy shock is now identified within a structural model rather than from a narrative series. Solving the RI-DSGE for the seven shocks simultaneously would be prohibitively expensive, ruling out full-information methods.[14] We can, however, proceed in two steps. We first estimate the NK-DSGE by full-information Bayesian methods, running a Metropolis–Hastings algorithm on the likelihood of seven quarterly aggregates — output, consumption, investment, hours, the real wage, inflation and the nominal interest rate; Appendix I provides the details. We then re-estimate the RI-DSGE with Equation (21), where $\widehat{\mathcal{J}}$ now collects the posterior medians of the NK-DSGE's impulse responses to a monetary policy shock, for the same four variables as in the direct estimation, and the weighting matrix is diagonal, carrying their posterior variances.

The models are compared by the overlap of their credible bands rather than by marginal data densities. The exercise answers a different question from the direct estimation: not whether the RI-DSGE fits the data as well as the NK-DSGE does, but whether it can reproduce the NK-DSGE at its posterior. Since we do not compute marginal data densities here, the prior does not affect the comparison; we leave the RI parameters unconstrained, as described in Appendix H.

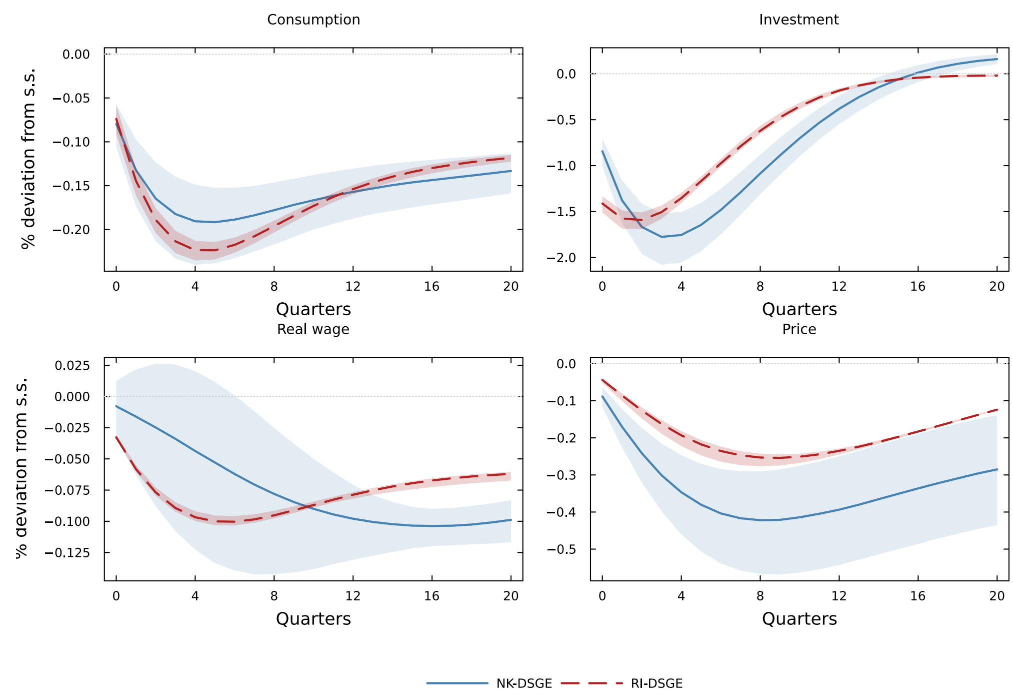
The NK parameters that enter the RI-DSGE need not be fixed at their posterior median: we treat their posterior as a prior and draw (ρ_m, σ_m) from it at every step of the Metropolis–Hastings algorithm, so that the quasi-posterior of $(\lambda_i, \lambda_j, \chi_k)$ reflects the uncertainty surrounding the monetary policy process itself. Thus, the process for the monetary policy shock block is itself updated by the estimation: its quasi-posterior need not coincide with the NK posterior it started from. In practice it barely moves — the quasi-posterior medians are $\rho_m = 0.048$ and $100\sigma_m = 0.214$, against 0.044 and 0.220 at the NK posterior.

The procedure is of interest beyond our application. When two models are estimated separately by full information, each infers its own exogenous processes, and differences in fit confound differences in propagation with differences in the estimated shock processes. Here the RI-DSGE inherits the monetary policy process estimated within the NK-DSGE and is asked to reproduce its transmission: whatever separates the two models is then attributable to the propagation mechanism alone.[15] The design is in the spirit of the sequential partial indirect inference of Dridi, Guay, and Renault (2007): the parameters of interest are matched on a deliberately restricted

[14]Under our specification of the attention problem, the complete model requires solving one fixed point per shock. Under the alternative specification, in which decision-makers attend to all shocks jointly, the state space becomes too large.

[15]Up to the movement in (ρ_m, σ_m) the estimation allows; one can instead hold the stochastic process fixed.

Figure 4: NK posterior impulse responses to a monetary policy shock vs. model fit



Note: Blue lines and shaded bands: posterior median responses and 90% credible bands of the NK-DSGE to a monetary policy shock. Red lines and bands: quasi-posterior median and 90% bands of the RI-DSGE estimated on the NK posterior medians.

set of instrumental characteristics, while the remaining structural parameters are held outside the matching step.

Figure 4 displays the result. The RI-DSGE reproduces the NK-DSGE's transmission, including the hump shapes of consumption and investment, at attention costs ($\lambda_i = 188$, $\lambda_j = 218$) and a capital adjustment cost of $\chi_k = 18.5$ (Table 1). There are two main discrepancies. The RI price level settles above the NK one, and investment dips deeper on impact before reverting faster.

Hybrid models. The same procedure allows for another exercise: holding one side of the economy at its New Keynesian frictions while inattention operates on the other. We define two hybrid models. In the RINK-DSGE, firms are inattentive and households face habit formation and wage rigidity; in the NKRI-DSGE, households are inattentive and firms face sticky prices and investment adjustment costs. Both are estimated exactly as the RI-DSGE is, with the NK parameters they retain drawn from their posterior alongside (ρ_m, σ_m). Comparing the two hybrids to the fully inattentive model isolates how much of the fit each side of the economy is responsible for. Figure 19 to Figure 24 summarize the results. Unsurprisingly, both hybrids reproduce the NK-DSGE posterior more closely than the fully inattentive model: on one side of the economy the propagation mechanism is the NK-DSGE's own, and the estimation only needs inattention to stand in for the frictions on the other. The comparison between the two is the informative part: conditional on a monetary policy shock, the NKRI-DSGE comes closer than the RINK-DSGE, indicating that inattentive households substitute for nominal wage rigidities and habits more easily than inattentive firms with capital adjustment costs substitute for nominal price rigidities and investment adjustment costs.

5.5. The Importance of Capital Adjustment Costs. The estimation puts a number on the mechanism of Section 3. Figure 5 re-solves the model at the indirect quasi-posterior median attention costs with the capital adjustment cost removed ($\chi_k = 0$), holding everything else fixed. Investment falls on impact by 18.8 percent of its steady-state level — thirteen times the estimated model's response — and every response is deepest in the first quarter: the hump shapes disappear along with the delay. Thus, without a friction penalizing adjustment of the capital stock, even slowly-revised estimates of the optimal target translate into large investment flows.

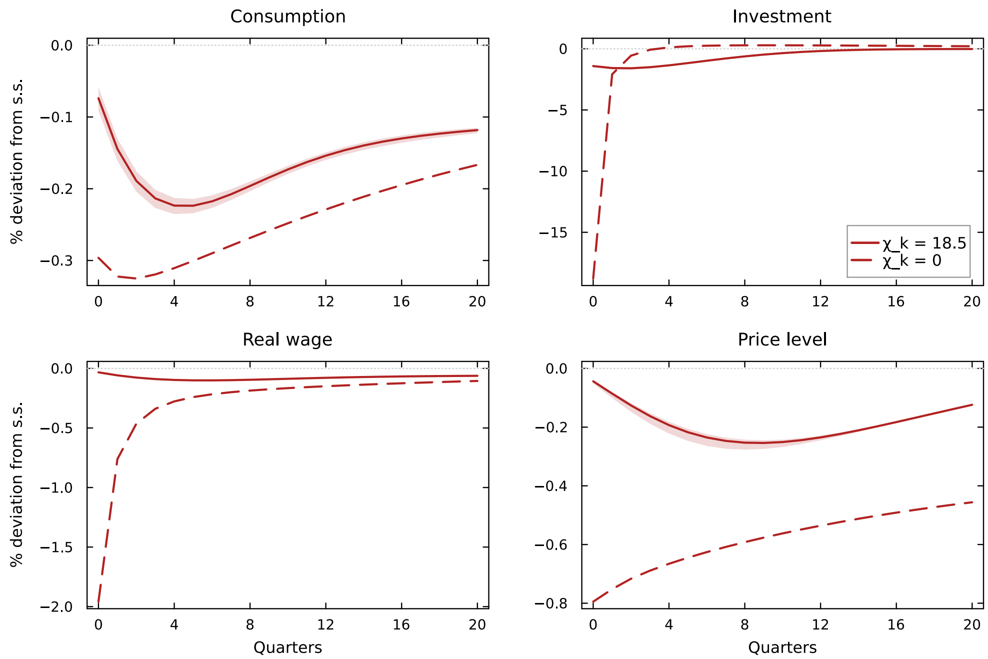
6. Beyond Monetary Policy Shocks

The comparison so far has rested on a single shock. The indirect estimation of Section 5.4 lifts that restriction: alongside the monetary policy shock process, the NK-DSGE posterior delivers six other business cycle shocks. Feeding them into the indirectly estimated RI-DSGE gives its impulse responses to every shock in the model. Those responses are the model's moving-average representation: they determine the covariance structure of the observables, from which a likelihood can be evaluated. We can therefore ask not only whether the two models agree about the transmission of monetary policy, but whether they agree about the business cycle, at the estimated processes of the NK model.

The question is worth asking because the two models need not coincide beyond the monetary policy shock: the RI-DSGE was estimated to reproduce the NK-DSGE's responses to that shock alone, through a different propagation mechanism, and nothing constrains the two mechanisms to deliver the same responses to the six untargeted shocks.

6.1. Likelihood. The model's moving-average representation directly delivers a likelihood, as in Mankiw and Reis (2007) and Auclert et al. (2021). Let $\boldsymbol{V}(\boldsymbol{\theta})$ denote

Figure 5: Impulse responses without capital adjustment costs



Note: Responses to a monetary policy shock. Solid lines: quasi-posterior median responses of the RI-DSGE at the indirect estimates ($\chi_k = 18.5$), with 90% credible bands. Dashed lines: responses re-solved at the quasi-posterior median attention costs with the capital adjustment cost removed ($\chi_k = 0$), all remaining parameters unchanged.

the covariance matrix of the n observables stacked over the sample, which the impulse responses to the seven shocks determine, and let $\boldsymbol{\mathcal{Y}}$ denote the data. Under Gaussian innovations,

$$\mathcal{L}(\boldsymbol{\mathcal{Y}};\boldsymbol{\theta}) = -\tfrac{1}{2}\log\det \boldsymbol{V}(\boldsymbol{\theta}) - \tfrac{1}{2}\boldsymbol{\mathcal{Y}}'\boldsymbol{V}(\boldsymbol{\theta})^{-1}\boldsymbol{\mathcal{Y}}. \tag{22}$$

Once the impulse responses are in hand, Equation (22) costs nothing further to evaluate, so combining draws from the RI quasi-posterior with draws from the NK posterior over the stochastic processes yields a posterior distribution of the RI-DSGE's log-likelihood. The comparison is conditional: the exogenous processes are those estimated to fit the NK-DSGE on the full covariance structure of aggregate time series. As a benchmark, we also re-estimate the NK-DSGE indirectly on its own posterior responses and evaluate it the same way, which separates what is lost to the two-step procedure from what is lost to the model. Recall that only the monetary policy responses entered the estimation; as long as the RI-DSGE matches the NK-DSGE's

Table 3: Posterior Distribution of the Log-Likelihood

	Mean	Median	[5^{th}, 95^{th}]
Full Information			
NK-DSGE	−871.9	−871.5	[−889.5, −855.7]
Indirect (Limited Information)			
NK-DSGE	−875.6	−874.1	[−894.4, −861.1]
RI-DSGE	−1990.1	−1991.7	[−2057.6, −1909.5]
RINK-DSGE	−1699.5	−1687.3	[−1984.1, −1468.6]
NKRI-DSGE	−3040.2	−2895.2	[−4334.2, −2276.7]

Notes: Log-likelihood Equation (22) of consumption, investment, the real wage and inflation over 1966Q1–2018Q4. In the indirect case, the NK-DSGE is re-estimated on the monetary policy shock IRFs generated at its own posterior, the same procedure used for the models with RI. The model's moving-average representation is truncated at the sample length

response to monetary policy, what remains is a loss on untargeted moments.[16]

Three features of Table 3 stand out. First, the two-step procedure itself is nearly costless: re-estimated indirectly on its own posterior responses, the NK-DSGE loses about 3 log points relative to its full-information fit. The remaining gaps are therefore attributable to the models, not to the method. Second, the RI-DSGE's median log-likelihood lies roughly 1,100 log points below the NK-DSGE's; we show in the next subsection that the gap loads almost entirely on the markup shocks. Third, the ranking of the hybrids reverses relative to Section 5.4: on the full set of shocks the RINK-DSGE likelihood comes closest to that of the NK-DSGE.

6.2. Impulse Response Functions. The large likelihood differences between the NK-DSGE and the models with rational inattention imply that their impulse responses must differ substantially: these responses are the model's moving-average representation, so they summarize the entire covariance structure the likelihood evaluates. The disagreement, however, need not be uniform across shocks — and it is not.

Figures 6 and 7 show that for the two shocks that account for most of the variation in quantities — TFP and MEI, see Figure 13 — the responses of the models with rational inattention are almost identical to those of the NK-DSGE. The RI-DSGE and its variants thus produce the right comovements for the shocks that Chari, Kehoe, and McGrattan (2009) view as structural; those that are interpretable.[17]

[16] We use the NK posterior for all seven processes, including monetary policy; the indirect estimation barely moves the latter.

[17] Justiniano, Primiceri, and Tambalotti (2011) find the MEI shock to be the most important

Figure 6: Impulse responses to a technology shock

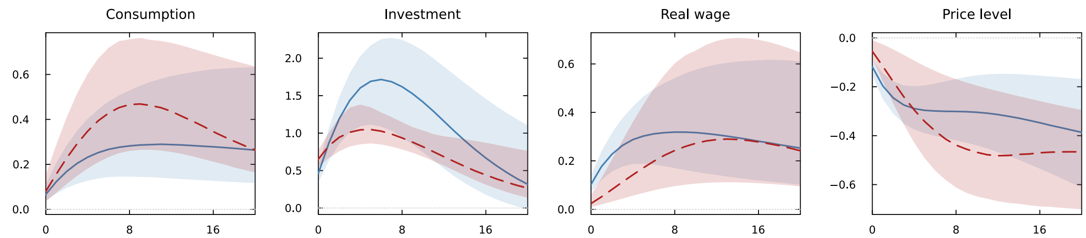
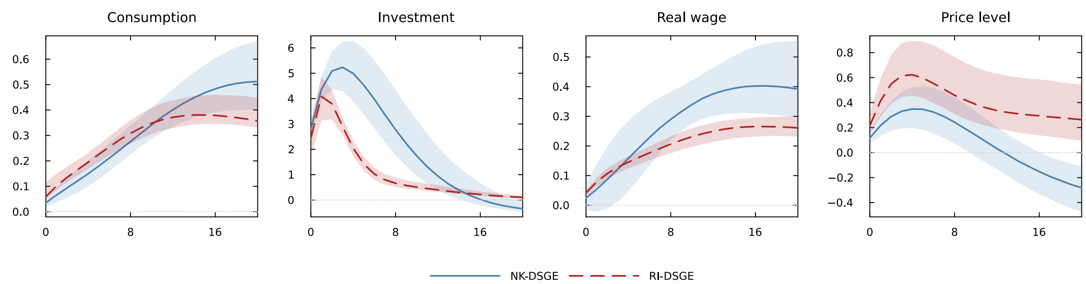


Figure 7: Impulse responses to an investment shock



Note: Solid line posterior median; shaded area 90% credible bands. RI-DSGE indirectly estimated on the NK-DSGE IRFs to a monetary policy shock.

Technology shocks have an empirical counterpart measured directly from the data, and the estimated process largely coincides with it (Fernald 2014); the MEI shock has no counterpart as direct, but its process travels across models that share the same core and match the same monetary policy transmission — which is what structural processes should do. The success here is not agreement with the NK-DSGE for its own sake: for these shocks, the NK responses are the empirically credible ones, and our model reproduces them with a single primitive in place of the adjustment frictions — estimated on the transmission of one shock, it delivers the transmission of the two untargeted shocks that drive the bulk of the fluctuations.

The responses that disagree the most, shown in Figures 8 and 9, are those to the markup shocks. Those shocks have long been criticized as "dubious" — not structural, but reduced-form wedges standing in for fluctuations the NK-DSGE is not built to explain (Chari, Kehoe, and McGrattan, 2007). That the RI-DSGE

driver of postwar U.S. business cycle fluctuations and interpret it as a proxy for disturbances to the functioning of the financial sector, a channel absent from our model; they corroborate the interpretation by showing that the estimated shock correlates strongly with interest rate spreads.

Figure 8: Impulse responses to a price markup shock

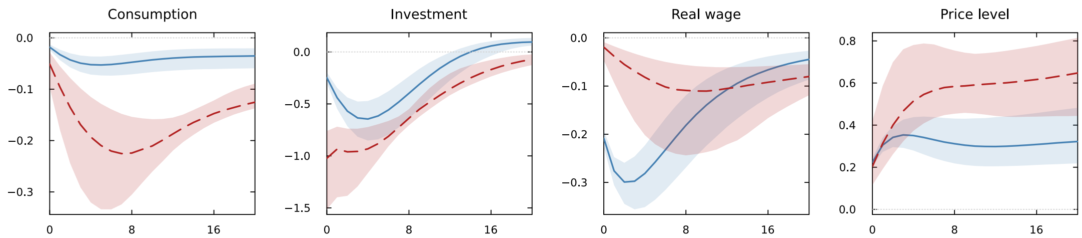
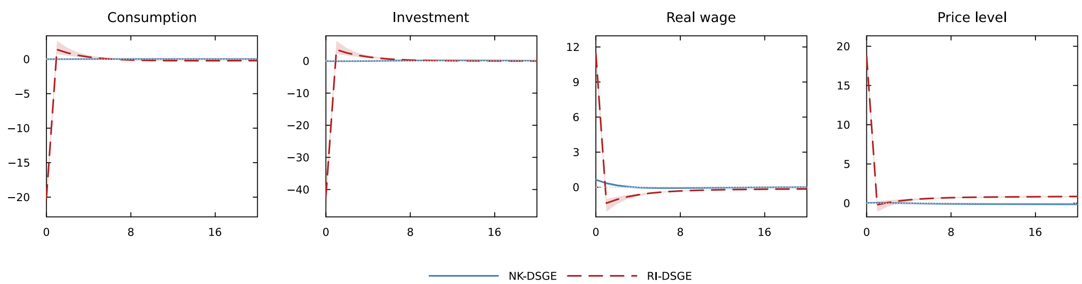


Figure 9: Impulse responses to a wage markup shock



Note: Solid line posterior median; shaded area 90% credible bands. RI-DSGE indirectly estimated on the NK-DSGE IRFs to a monetary policy shock.

departs from the NK-DSGE precisely there is unsurprising. The processes of the markup shocks are estimated given the NK propagation mechanism, and are tailored to absorb what the rest of that model leaves unexplained; the RI-DSGE has no reason to reproduce the same dynamics. One could estimate markup processes that raise the RI-DSGE's likelihood, but they would bear no relation to those the NK-DSGE recovers — which only cements the point that these shocks are not structural. The RI-DSGE rejects precisely the shocks whose structural interpretation is in doubt.

Taken together, the exercise sharpens the comparison between the two models. At the shock processes the NK-DSGE itself recovers, the two models describe the same business cycle wherever the shocks admit a structural interpretation, and part ways only on the wedges — shocks that contribute little to our understanding of business cycles and rather mark its limits. The conjecture of Sims thus extends beyond the monetary policy shock.

7. Key Differences Between the Models

This section presents additional differences — beyond the dynamic responses to macro shocks — between the models. We consider two exercises: the firm-level responses to idiosyncratic shocks, and a counterfactual regime in which the central bank's response to inflation weakens.

7.1. Firm Responses to Idiosyncratic Shocks. The fact that the RI-DSGE produces aggregate dynamics similar to those of the NK-DSGE wherever the shocks admit a structural interpretation — while the latter relies on a combination of ad-hoc adjustment frictions — is in itself a reason to rethink the sources of inertia in macro models. In this subsection, we cement this point by comparing the firm-level responses of prices and investment to idiosyncratic shocks in the two models. Empirical studies document that these shocks induce monotonic responses peaked on impact (see Boivin, Giannoni, and Mihov 2009 for prices and Zorn 2020 for investment); appropriate microfoundations should reproduce this pattern — indeed, reproducing micro evidence is the very criterion Altig et al. (2011) propose for adjudicating among candidates: "the debate will be settled on the field of microeconomic data" (p. 228).

Idiosyncratic and aggregate shocks generate qualitatively different responses for two standard reasons. The first applies equally to both models: strategic complementarities and general equilibrium feedbacks dampen responses to aggregate conditions, and at the firm level, where aggregates are held fixed, both are absent by construction. The second is where the models part ways: firm-level shocks are larger and more volatile than aggregate ones, and only under rational inattention does that difference matter. The NK frictions dampen responses regardless of the shock process; the NK firm's decision rules are thus invariant to it. Such invariance is exactly what the *Lucas critique* warns against; we apply the critique to idiosyncratic shocks here, but the logic extends to changes in the aggregate environment as well. Under RI, firms' attention is endogenous, so the same λ_i that delivers sluggish responses to aggregate shocks need not imply sluggish responses to volatile idiosyncratic ones.

We solve each model's firm-level problem at its own estimated frictions — the posterior medians of (κ_p, χ_i) for the NK firm (Appendix E.1) and of (λ_i, χ_k) for the RI firm (Appendix A.6) — under two idiosyncratic technology processes, each the quarterly equivalent of its source literature. The NK firm faces Rotemberg pricing frictions, so that both firms adjust smoothly and their responses are directly comparable. The first process is the sectoral one of Zorn (2020): quarterly persistence 0.86 and an unconditional standard deviation roughly four times that of the aggregate technology process. The second is the firm-level process of Maćkowiak and Wieder-

Figure 10: Maćkowiak and Wiederholt (2015)'s process

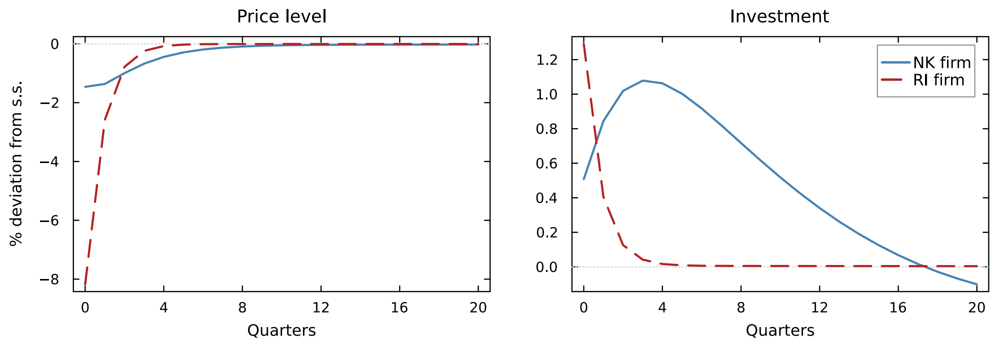
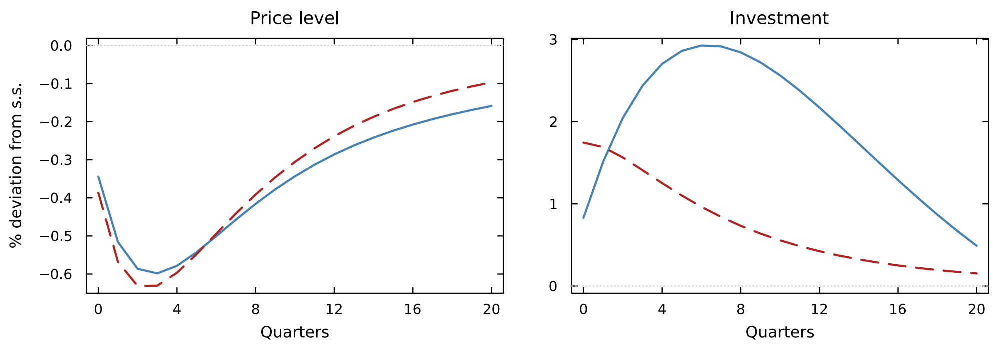


Figure 11: Zorn (2020)'s process



Note: RI-DSGE indirectly estimated on the NK-DSGE IRFs to a monetary policy shock.

holt (2015), chosen to match the average absolute size of price changes observed in micro data: persistence 0.30 and standard deviation tenfold that of aggregate technology.

Figures 10 and 11 show that the RI firm reproduces the empirical pattern under both processes. Its investment response peaks on impact and declines monotonically, matching the sectoral evidence in Zorn (2020). Its price response adjusts by about a third of the frictionless adjustment on impact under the sectoral process, and by over 90 percent under the firm-level process — the near-perfect tracking of firm-specific conditions emphasized by Maćkowiak and Wiederholt (2015) and consistent with the fast sectoral price responses in Boivin, Giannoni, and Mihov (2009). The NK firm, by contrast, fails under both specifications. Its investment response is hump-shaped — peaking five to six quarters after a sectoral shock at roughly three times its impact value, and still hump-shaped in year one even under the nearly transitory firm-level process — and its price adjusts by the same small fraction of the frictionless response regardless of the properties of the shock.

7.2. A Taylor Rule with a Weak Response to Inflation. *(TBC)*

8. Conclusions

We construct, solve and estimate a medium-scale DSGE model in which firms and households are rationally inattentive. We show that limited attention and strategic complementarities alone cannot account for the inertia and persistence of aggregate quantities: in general equilibrium, the slow adjustment of prices amplifies the optimal investment response, and a real friction on capital accumulation is required. The finding establishes a limitation of inattention in general equilibrium, so far documented only in partial-equilibrium settings.

Estimated on the impulse responses to a monetary policy shock, our model fits the evidence as well as a prototypical NK-DSGE, with the marginal data densities even favoring the model with rational inattention. Fed the NK-estimated shock processes, our model implies transmission of the technology and investment-efficiency shocks that drive most of the business cycle that are alike; the likelihood gap between the two models loads almost entirely on the markup shocks, long criticized as reduced-form wedges. At the firm level, our model matches the fast, impact-peaked responses to idiosyncratic shocks documented in micro data, where the NK firm does not.

Taken together, our results support Sims' conjecture: the adjustment frictions of estimated New Keynesian models are well read as proxies for limited attention — with the exception of a real friction slowing down the response of investment, which captures a genuine rigidity in capital accumulation. Future work should seek a deeper microfoundation for that rigidity, in the spirit of time to build (Kydland and Prescott 1982). Our framework also admits several natural extensions: (i) letting decision-makers bundle information across some shocks, whose quantitative importance remains to be investigated, and (ii) letting households trade in the mutual fund. Lastly, moving away from markup shocks to account for the variation in prices seems necessary.

References

Afrouzi, Hassan, and Choongryul Yang. 2021. "Dynamic Rational Inattention and the Phillips Curve." CESifo Working Paper 8840, CESifo.

Altig, David, Lawrence Christiano, Martin Eichenbaum, and Jesper Lindé. 2011. "Firm-specific capital, nominal rigidities and the business cycle." *Review of Economic Dynamics* 14 (2): 225–247.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models." *Econometrica* 89 (5): 2375–2408.

Boivin, Jean, Marc P. Giannoni, and Ilian Mihov. 2009. "Sticky prices and monetary policy: Evidence from disaggregated US data." *American Economic Review* 99 (1): 350–384.

Chari, Varadarajan V, Patrick J Kehoe, and Ellen R McGrattan. 2007. "Business cycle accounting." *Econometrica* 75 (3): 781–836.

Chari, Varadarajan V, Patrick J Kehoe, and Ellen R McGrattan. 2009. "New Keynesian models: Not yet useful for policy analysis." *American Economic Journal: Macroeconomics* 1 (1): 242–266.

Chernozhukov, Victor, and Han Hong. 2003. "An MCMC approach to classical estimation." *Journal of Econometrics* 115 (2): 293–346.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of Political Economy* 113 (1): 1–45.

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt. 2016. "Unemployment and Business Cycles." *Econometrica* 84 (4): 1523–1569.

Christiano, Lawrence J., Mathias Trabandt, and Karl Walentin. 2010. "Involuntary Unemployment and the Business Cycle." NBER Working Paper 15801, National Bureau of Economic Research.

Dridi, Ramdan, Alain Guay, and Eric Renault. 2007. "Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models." *Journal of Econometrics* 136 (2): 397–430.

Fernald, John. 2014. “A Quarterly, Utilization-Adjusted Series on Total Factor Productivity.” Technical report, Federal Reserve Bank of San Francisco.

Han, Zhao, Fei Tan, and Jieran Wu. 2022. “Analytic policy function iteration.” *Journal of Economic Theory* 200: 105395.

Hayashi, Fumio. 1982. “Tobin’s Marginal q and Average q: A Neoclassical Interpretation.” *Econometrica* 50 (1): 213–224.

Inoue, Atsushi, and Mototsugu Shintani. 2018. “Quasi-Bayesian Model Selection.” *Quantitative Economics* 9 (3): 1265–1297.

Jordà, Òscar. 2005. “Estimation and inference of impulse responses by local projections.” *American Economic Review* 95 (1): 161–182.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2011. “Investment Shocks and the Relative Price of Investment.” *Review of Economic Dynamics* 14 (1): 102–121.

Kydland, Finn E., and Edward C. Prescott. 1982. “Time to Build and Aggregate Fluctuations.” *Econometrica* 50 (6): 1345–1370.

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

Maćkowiak, Bartosz, and Mirko Wiederholt. 2015. “Business cycle dynamics under rational inattention.” *The Review of Economic Studies* 82 (4): 1502–1532.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2025. “Rational Inattention and the Business Cycle Effects of Productivity and News Shocks.” *American Economic Journal: Macroeconomics* 17 (1): 274–309.

Mankiw, N 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.

Mankiw, N. Gregory, and Ricardo Reis. 2007. “Sticky information in general equilibrium.” *Journal of the European Economic Association* 5 (2-3): 603–613.

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt. 2018. “Dynamic rational inattention: Analytical results.” *Journal of Economic Theory* 176: 650–692.

Miao, Jianjun, Jieran Wu, and Eric R Young. 2022. "Multivariate rational inattention." *Econometrica* 90 (2): 907–945.

Montiel Olea, José Luis, and Mikkel Plagborg-Møller. 2021. "Local Projection Inference Is Simpler and More Robust Than You Think." *Econometrica* 89 (4): 1789–1823.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708.

Nimark, Kristoffer. 2008. "Dynamic pricing and imperfect common knowledge." *Journal of Monetary Economics* 55 (2): 365–382.

Ramey, Valerie A. 2016. "Macroeconomic shocks and their propagation." In *Handbook of Macroeconomics*, vol. 2, 71–162: Elsevier.

Romer, Christina D, and David H Romer. 2004. "A new measure of monetary shocks: Derivation and implications." *American Economic Review* 94 (4): 1055–1084.

Rotemberg, Julio J., and Michael Woodford. 1997. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual* 12: 297–346.

Sims, Christopher A. 1998. "Stickiness." In *Carnegie-Rochester Conference Series on Public Policy*, vol. 49: 317–356, Elsevier.

Sims, Christopher A. 2002. "Solving Linear Rational Expectations Models." *Computational Economics* 20 (1): 1–20.

Smets, Frank, and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review* 97 (3): 586–606.

Song, Wenting, and Samuel Stern. 2025. "Firm Inattention and the Efficacy of Monetary Policy: A Text-Based Approach." *The Review of Economic Studies* 92 (5): 3438–3469.

Sveen, Tommy, and Lutz Weinke. 2003. "Inflation and Output Dynamics with Firm-Owned Capital." Economics Working Paper 702, Universitat Pompeu Fabra.

Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford. Princeton, NJ: Princeton University Press.

Zorn, Peter. 2020. "Investment under rational inattention: Evidence from US sectoral data." CESifo Working Paper 8436, CESifo.

A. RI-DSGE: Derivation of the Optimality Conditions

In this section, we derive the optimality conditions for firms and households with perfect information in the RI-DSGE. These conditions characterize the optimal actions in the attention problems.

Definition 1. *In any period t, a decision-maker with perfect information has rational expectations and knows the entire history of the economy up to and including current period t.*

A.1. Households. Every household j solves

$$\max_{\{C_{jt},\bar{L}_{jt},W_{jt},B_{jt},Q_{jt}\}_{t=0}^{\infty}} E_{jt}\left\{\sum_{t=0}^{+\infty}\beta^t \exp\{\epsilon_t^c\}\left(\frac{C_{jt}^{1-\gamma}-1}{1-\gamma}-\varphi\frac{\bar{L}_{jt}^{1+\psi}}{1+\psi}\right)\right\} \tag{23}$$

subject to Equation (2) and the labor demand schedule. The first-order conditions yield the Euler equation

$$\exp\{\epsilon_t^c\}=\beta E_{jt}\left[\exp\{\epsilon_{t+1}^c\}\left(\frac{C_{j,t+1}}{C_{jt}}\right)^{-\gamma}\left(\frac{R_t}{\Pi_{t+1}}\right)\right], \tag{24}$$

the no-arbitrage condition between bonds and shares

$$S_t=E_{jt}\left[(S_{t+1}+D_{t+1})\left(\frac{\Pi_{t+1}}{R_t}\right)\right], \tag{25}$$

and the optimal wage condition

$$\exp\{\psi\epsilon_{jt}^z\}W_{jt}=\varphi\frac{(1+\epsilon_t^w)}{(1+\tau_w)}\left(\hat{W}_{jt}^{-\vartheta_w}L_{jt}\right)^{\psi}C_{jt}^{\gamma}. \tag{26}$$

Equation (24) is the standard Euler equation linking current consumption to expected consumption next period and the real rate. Equation (25) is a no-arbitrage condition between bond holdings and shares in the mutual fund. Equation (26) sets the real wage as a markup over the marginal rate of substitution between consumption and labor supply.

A.2. Firms. Every firm i solves

$$\max_{\{P_{it},L_{it},K_{it},I_{it}\}_{t=0}^{\infty}} E_{it}\left\{\sum_{t=0}^{+\infty}Q_{-1,t}\left[(1+\tau_p)\hat{P}_{it}Y_{it}-W_tL_{it}-I_{it}-\tfrac{\chi_k}{2}\left(\tfrac{K_{it}}{K_{it-1}}-1\right)^2K_{it-1}\right]\right\} \tag{27}$$

subject to Equation (4), Equation (7), and Equation (5). The first-order conditions yield the optimal pricing equation

$$\hat{P}_{it} = \left(\frac{W_t}{\alpha}\right)\frac{(1+\epsilon_t^p)}{(1+\tau_p)}\left(\frac{Y_t^{1-\alpha}\hat{P}_{it}^{(\alpha-1)\vartheta_t^p}}{\exp\{\epsilon_t^a+\epsilon_{it}^a\}K_{it-1}^{\phi}}\right)^{\frac{1}{\alpha}} \tag{28}$$

and the investment Euler equation

$$\begin{aligned}
&\Lambda(\{C_{jt}\})\exp\{\epsilon_t^c\}\left[\frac{1}{\exp\{\epsilon_t^i\}}+\chi_k\left(\frac{K_{it}}{K_{it-1}}-1\right)\right] \\
&= \beta E_{it}\Bigg\{\Lambda(\{C_{jt+1}\})\exp\{\epsilon_{t+1}^c\}\Bigg[\frac{(1-\delta)}{\exp\{\epsilon_{t+1}^i\}}+\frac{\phi}{\alpha}W_{t+1}\left(\frac{Y_{t+1}\hat{P}_{it+1}^{-\vartheta_t^p}}{\exp\{\epsilon_{t+1}^a+\epsilon_{it+1}^a\}K_{it}^{\phi+\alpha}}\right)^{\frac{1}{\alpha}} \\
&\qquad +\chi_k\frac{K_{it+1}}{K_{it}}\left(\frac{K_{it+1}}{K_{it}}-1\right)-\frac{\chi_k}{2}\left(\frac{K_{it+1}}{K_{it}}-1\right)^2\Bigg]\Bigg\}.
\end{aligned} \tag{29}$$

Equation (28) is the standard pricing equation under monopolistic competition: the firm sets its price as a markup over marginal cost, the wage rate divided by the marginal product of labor. Equation (29) is an Euler equation for investment, equating the marginal cost of an additional unit of capital today to its expected discounted marginal contribution to the firm's value next period. On the left, that cost is the price of the investment good plus the adjustment cost the extra unit triggers. On the right, the contribution has three parts: the savings in labor cost from the extra unit of capital — marginal cost times the marginal product of capital — its post-depreciation resale value, and the reduction in next period's adjustment cost that a larger inherited stock brings. Setting $\chi_k = 0$ removes the first and last of these and recovers the frictionless Euler equation.

A.3. Non-Stochastic Steady State. The non-stochastic steady state satisfies Equation (30)–Equation (33):

$$W = \varphi L^{\psi} C^{\gamma} \tag{30}$$

$$\frac{R}{\Pi} = \frac{1}{\beta} \tag{31}$$

$$1 = \frac{W}{\alpha}\left(\frac{Y^{1-\alpha}}{K^{\phi}}\right)^{\frac{1}{\alpha}} \tag{32}$$

$$1 = \beta\left[(1-\delta)+\frac{\phi}{\alpha}W\left(\frac{Y}{K^{\phi+\alpha}}\right)^{\frac{1}{\alpha}}\right]. \tag{33}$$

The useful relationships $\alpha Y = WL$, $K = \beta\phi Y/[1-\beta(1-\delta)]$, $\varphi L^{1+\psi} = C^{1-\gamma}\omega_W$, and $(1-\beta)\omega_S = \beta\omega_D$ follow directly, where $\omega_W := WL/C$, $\omega_S := S/C$, $\omega_D := D/C$.

A.4. Linearized Optimality Conditions. Dropping idiosyncratic shocks (they sum to zero in the cross-section) and log-linearizing:

$$c_{jt} = E_{jt}[c_{jt+1}] - \tfrac{1}{\gamma} E_{jt}[r_t - \pi_{t+1} + (\rho_c - 1)\epsilon_t^c] \tag{34}$$

$$w_{jt} = \tfrac{1}{1+\psi\vartheta^w}\left[\psi l_t + \psi\vartheta^w w_t + \gamma c_{jt} + \hat{\epsilon}_t^w - \psi\epsilon_{jt}^z - \epsilon_t^c\right] \tag{35}$$

$$p_{it} = p_t + \left(1 + \tfrac{(1-\alpha)\vartheta^p}{\alpha}\right)^{-1}\left[w_t + \tfrac{1-\alpha}{\alpha} y_t - \tfrac{1}{\alpha}(\epsilon_t^a + \epsilon_{it}^a) - \tfrac{\phi}{\alpha} k_{it-1} + \hat{\epsilon}_t^p\right] \tag{36}$$

$$\begin{aligned} \chi_k\big[(k_{it} - k_{it-1}) - \beta E_{it}(k_{it+1} - k_{it})\big] = \\ E_{it}\Big[\gamma(c_t - c_{t+1}) + (\rho_c - 1)\epsilon_t^c + \big(1 - \rho_i\beta(1-\delta)\big)\epsilon_t^i \\ + \big(1 - \beta(1-\delta)\big)\Big(w_{t+1} + \tfrac{1}{\alpha} y_{t+1} - \tfrac{1}{\alpha}(\epsilon_{t+1}^a + \epsilon_{it+1}^a) - \tfrac{\alpha+\phi}{\alpha} k_{it}\Big)\Big] \end{aligned} \tag{37}$$

where $\hat{\epsilon}_t^p := \log(1+\epsilon_t^p) - \log(1+\epsilon^p)$ and $\hat{\epsilon}_t^w := \log(1+\epsilon_t^w) - \log(1+\epsilon^w)$.

A.5. Remaining Equilibrium Conditions. The conditions above determine individual actions. Aggregation, market clearing, technology and policy close the model:

$$p_t = \textstyle\int_0^1 p_{it}\, di \tag{38}$$

$$w_t = \textstyle\int_0^1 w_{jt}\, dj \tag{39}$$

$$c_t = \textstyle\int_0^1 c_{jt}\, dj \tag{40}$$

$$k_t = \textstyle\int_0^1 k_{it}\, di \tag{41}$$

$$\pi_t = p_t - p_{t-1} \tag{42}$$

$$y_t = a_t + \alpha l_t + \phi k_{t-1} \tag{43}$$

$$i_t = \tfrac{1}{\delta} k_t - \tfrac{1-\delta}{\delta} k_{t-1} - \epsilon_t^i \tag{44}$$

$$y_t = \tfrac{C}{Y} c_t + \tfrac{I}{Y} i_t + \tfrac{G}{Y}\epsilon_t^g \tag{45}$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\phi_\pi \pi_t + (1-\rho_r)\phi_{y^*} y_t^* + \epsilon_t^m \tag{46}$$

where y_t^* is output in the economy without information frictions and markup shocks.

A.6. Firms' Equilibrium Conditions to an Idiosyncratic Shock. The following conditions determine the profit-maximizing actions of a rationally inattentive firm i subject to capital adjustment costs, in response to an idiosyncratic productivity shock ϵ^a_{it}. Aggregate variables are unaffected.

$$p^*_{it} = mc_{it} \tag{47}$$

$$mc_{it} = (1-\alpha)l_{it} - \phi k_{it-1} - \epsilon^a_{it} \tag{48}$$

$$\chi_k\left[(k_{it}-k_{it-1}) - \beta E_{it}(k_{it+1}-k_{it})\right] = \big(1-\beta(1-\delta)\big)\, E_{it}[mpk_{it+1} + mc_{it+1}] \tag{49}$$

$$mpk_{it} = \epsilon^a_{it} + (\phi-1)k_{it-1} + \alpha l_{it} \tag{50}$$

$$-\vartheta^p p_{it} = \epsilon^a_{it} + \phi k_{it-1} + \alpha l_{it} \tag{51}$$

$$i_{it} = \tfrac{1}{\delta}k_{it} - \tfrac{1-\delta}{\delta}k_{it-1} \tag{52}$$

Solved under perfect information, these conditions define the processes for $\{p^*_{it}, k^*_{it}\}$. We solve the rational inattention problem with the same iterative procedure as in Appendix D, except that the only object we iterate on is the stochastic sequence for the firm's labor input l_{it}.

B. Derivation of the Firms' Losses from Suboptimal Actions

We approximate firm i's expected discounted real profits with a second-order log Taylor expansion around the non-stochastic steady state. Multiplying period profits by households' marginal utility and substituting the demand function, production function, and capital law of motion yields

$$\begin{aligned} f(\hat{p}_{it}, k_{it}, k_{it-1}, w_t, y_t, c_t, \epsilon^c_t, \epsilon^a_t, \epsilon^a_{it}, \epsilon^i_t, \epsilon^p_t) = \Lambda\big(\{Ce^{c_t}\}\big)\exp\{\epsilon^c_t\}\, Y \Bigg[& (1+\tau_p)\exp\{(1-\vartheta^p_t)\hat{p}_{it} + y_t\} \\ & - \alpha \exp\Big\{\tfrac{w_t + y_t - \vartheta^p_t \hat{p}_{it} - \phi k_{it-1} - \epsilon^a_t - \epsilon^a_{it}}{\alpha}\Big\} \\ & - \tfrac{\beta\phi}{1-\beta(1-\delta)} \cdot \tfrac{\exp\{k_{it}\} - (1-\delta)\exp\{k_{it-1}\}}{\exp\{\epsilon^i_t\}} \\ & - \tfrac{\chi_k}{2}\left(\exp\{k_{it}-k_{it-1}\} - 1\right)^2 \exp\{k_{it-1}\} \Bigg]. \end{aligned} \tag{53}$$

We used the non-stochastic steady-state relationships from Appendix A.3 to replace the wage bill WL and capital K. Let $F(\cdot)$ denote the expected discounted sum of $f(\cdot)$, obtained by multiplying the latter by the appropriate time discount factor β^t and summing over all t from zero to infinity, and let $\tilde{F}(\cdot)$ denote its second-order Taylor approximation around the non-stochastic steady state.

Proposition 1. *Let $\boldsymbol{x}_{it} = (p_{it}, k_{it})'$ denote the firm's vector of actions, $\boldsymbol{z}_{it}$ the variables taken as given, and $\boldsymbol{v}_{it} = (\boldsymbol{x}'_{it}, \boldsymbol{z}'_{it}, 1)'$, with v_{imt} denoting the m-th element of $\boldsymbol{v}_{it}$; since the aggregate price level is taken as given, $p_{it} - p^*_{it} = \hat{p}_{it} - \hat{p}^*_{it}$ and the losses* *below are unchanged if the firm's action is written as a relative price. Under condition $E_{i,-1}|v_{imt}v_{int+\tau}| < \varsigma^t A$ for some $\varsigma < \beta^{-1}$ and $A \in \mathbb{R}$, the expected discounted sum of losses in real profits is*

$$\begin{aligned} & E_{i,-1}\left[\tilde{F}(k_{i-1}, \boldsymbol{x}_{i0}, \boldsymbol{z}_{i0}, \cdots)\right] - E_{i,-1}\left[\tilde{F}(k_{i-1}, \boldsymbol{x}^*_{i0}, \boldsymbol{z}_{i0}, \cdots)\right] \\ = & \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left[\frac{1}{2}(\boldsymbol{x}_{it} - \boldsymbol{x}^*_{it})'\boldsymbol{\Theta}_0(\boldsymbol{x}_{it} - \boldsymbol{x}^*_{it}) + (\boldsymbol{x}_{it} - \boldsymbol{x}^*_{it})'\boldsymbol{\Theta}_1(\boldsymbol{x}_{it+1} - \boldsymbol{x}^*_{it+1})\right] \end{aligned} \tag{54}$$

where

$$\boldsymbol{\Theta}_0 = -C^{-\gamma}Y \begin{bmatrix} \Theta_0^p & 0 \\ 0 & \Theta_0^k \end{bmatrix}, \qquad \boldsymbol{\Theta}_1 = -C^{-\gamma}Y \begin{bmatrix} 0 & 0 \\ \Theta_1^{kp} & \Theta_1^k \end{bmatrix}, \tag{55}$$

with

$$\Theta_0^p = \vartheta^p\left(1 + \tfrac{1-\alpha}{\alpha}\vartheta^p\right), \qquad \Theta_0^k = \frac{\beta\phi(\alpha+\phi)}{\alpha} + \frac{\beta\phi(1+\beta)\chi_k}{1-\beta(1-\delta)}, \tag{56}$$

$$\Theta_1^{kp} = \frac{\beta\phi\vartheta^p}{\alpha}, \qquad \Theta_1^k = -\frac{\beta^2\chi_k\phi}{1-\beta(1-\delta)}. \tag{57}$$

When $\chi_k = 0$, $\Theta_1^k = 0$ and Θ_0^k reduces to $\beta\phi(\alpha+\phi)/\alpha$.

Proof. Take a second-order Taylor expansion of $F(\cdot)$ around the non-stochastic steady state, then use the firms' first-order conditions to eliminate terms involving variables taken as given, leaving only expressions in $(\boldsymbol{x}_{it} - \boldsymbol{x}^*_{it})$. These steps are identical to steps "Second" to "Seventh" in Maćkowiak and Wiederholt (2015), Appendix D, which apply whenever the period payoff depends only on current actions $\boldsymbol{x}_{it}$, past actions $\boldsymbol{x}_{it-1}$, and variables taken as given $\boldsymbol{z}_{it}$ — a structure shared by $f(\cdot)$. The Hessians are obtained by differentiating $f(\cdot)$ twice with respect to $\boldsymbol{x}_{it}$, which gives $\boldsymbol{\Theta}_0$, and with respect to $(\boldsymbol{x}_{it-1}, \boldsymbol{x}_{it})$, which gives $\boldsymbol{\Theta}_1$ once the date-$(t+1)$ payoff is reindexed and its extra discount factor β carried along, both evaluated at the non-stochastic steady state. □

The loss Equation (54) is not yet a tracking problem: the firm's current mistakes interact with its future ones through $\boldsymbol{\Theta}_1$. Consider first the case without capital adjustment costs, $\chi_k = 0$, where the only such interaction couples today's investment mistake with tomorrow's pricing mistake.

Proposition 2. *Let $\chi_k = 0$ and define*

$$\zeta_{it} := p_{it} + \tfrac{\phi}{\vartheta^p + \alpha(1-\vartheta^p)}k_{it-1}, \qquad \zeta^*_{it} := p^*_{it} + \tfrac{\phi}{\vartheta^p + \alpha(1-\vartheta^p)}k^*_{it-1}. \tag{58}$$

Under the same condition as Proposition 1, Equation (54) equals

$$-C^{-\gamma}Y\sum_{t=0}^{\infty}\beta^t E_{i,-1}\Big[\tfrac{1}{2}\Big(\Theta_0^p(\zeta_{it}-\zeta_{it}^*)^2+\bar{\Theta}_0^k\,(k_{it}-k_{it}^*)^2\Big)\Big], \tag{59}$$

with

$$\bar{\Theta}_0^k := \frac{\beta\phi[\vartheta^p(1-\alpha-\phi)+(\alpha+\phi)]}{\vartheta^p+\alpha(1-\vartheta^p)}. \tag{60}$$

Proof. Substituting the definition of ζ_{it} gives the identity $(p_{it}-p_{it}^*)=(\zeta_{it}-\zeta_{it}^*)-\frac{\phi}{\vartheta^p+\alpha(1-\vartheta^p)}(k_{it-1}-k_{it-1}^*)$. Multiply by β^t, sum from $t=0$ to T, use $k_{i,-1}^*=k_{i,-1}$, and telescope. Take expectations $E_{i,-1}$ and the limit $T\to\infty$. □

The change of variables is analogous to Maćkowiak and Wiederholt (2025), where firms also have firm-specific capital but choose labor rather than prices. Absent capital adjustment costs it delivers a diagonal tracking problem directly: the firm tracks ζ_{it}^* and k_{it}^*, and the coupling has been absorbed into the redefined price target. The curvature of pricing losses is unaffected, which is why Θ_0^p reappears in Equation (59); only that of capital is modified.

Capital adjustment costs. With $\chi_k>0$ the argument is incomplete. Two entries of Equation (55) that vanish at $\chi_k=0$ are now nonzero: the within-period curvature of capital carries the term $\beta\phi(1+\beta)\chi_k/[1-\beta(1-\delta)]$, which simply adds to $\bar{\Theta}_0^k$, and Θ_1^k couples today's capital mistake with tomorrow's. The ζ substitution telescopes the price–capital coupling only and leaves the second one untouched, so Proposition 2 now reads

$$\begin{aligned}-C^{-\gamma}Y\sum_{t=0}^{\infty}\beta^t E_{i,-1}\Big[&\tfrac{1}{2}\Big(\Theta_0^p(\zeta_{it}-\zeta_{it}^*)^2+\bar{\Theta}_0^k\,(k_{it}-k_{it}^*)^2\Big)\\ &+\Theta_1^k\,(k_{it}-k_{it}^*)(k_{it+1}-k_{it+1}^*)\Big],\end{aligned} \tag{61}$$

with $\bar{\Theta}_0^k$ now including the χ_k term,

$$\bar{\Theta}_0^k = \frac{\beta\phi[\vartheta^p(1-\alpha-\phi)+(\alpha+\phi)]}{\vartheta^p+\alpha(1-\vartheta^p)}+\frac{\beta\phi(1+\beta)\chi_k}{1-\beta(1-\delta)}. \tag{62}$$

A second change of variables removes the remaining coupling.

Proposition 3. *Let $\chi_k>0$ and let ν^* be the root of*

$$\Theta_1^k\,\nu^2-\bar{\Theta}_0^k\,\beta\,\nu+\Theta_1^k\,\beta=0 \tag{63}$$

lying outside the unit circle. Define

$$\Delta\tilde{k}_{it} := k_{it-1} + \nu^*\, k_{it}, \qquad \Delta\tilde{k}^*_{it} := k^*_{it-1} + \nu^* k^*_{it}, \tag{64}$$

and $\tilde{\Theta}^k_0 := \Theta^k_1/(\nu^*\beta)$. *Under the same condition as Proposition 1, Equation (54) equals*

$$\mathcal{L} = -C^{-\gamma} Y \sum_{t=0}^{\infty} \beta^t E_{i,-1}\Big[\tfrac{1}{2}\Big(\Theta^p_0\,(\zeta_{it} - \zeta^*_{it})^2 + \tilde{\Theta}^k_0\,(\Delta\tilde{k}_{it} - \Delta\tilde{k}^*_{it})^2\Big)\Big]. \tag{65}$$

Proof. Write the capital block of Equation (61) as $\frac{1}{2}\tilde{\Theta}^k_0 \sum_{t=0}^{\infty} \beta^t(\Delta\tilde{k}_{it} - \Delta\tilde{k}^*_{it})^2$ for some scalars ν and $\tilde{\Theta}^k_0$. Expanding and applying the index shifts $\sum_{t=0}^{\infty} \beta^t(k_{it-1} - k^*_{it-1})^2 = \beta\sum_{t=0}^{\infty} \beta^t(k_{it} - k^*_{it})^2$ and $\sum_{t=0}^{\infty} \beta^t(k_{it-1} - k^*_{it-1})(k_{it} - k^*_{it}) = \beta\sum_{t=0}^{\infty} \beta^t(k_{it} - k^*_{it})(k_{it+1} - k^*_{it+1})$, the coefficients on $(k_{it} - k^*_{it})^2$ and on the cross-product match provided $\tilde{\Theta}^k_0 = \Theta^k_1/(\nu\beta)$ and ν solves Equation (63).

Since $\Theta^k_1 < 0$, Equation (63) can be written $\nu^2 + b\nu + \beta = 0$ with $b = \bar{\Theta}^k_0\beta/|\Theta^k_1| > 1 + \beta$. Its roots are real and negative, their product is β and their absolute values sum to more than $1 + \beta$, so exactly one lies outside the unit circle. Selecting it makes the recursion $(k_{it} - k^*_{it}) = (\Delta\tilde{k}_{it} - \Delta\tilde{k}^*_{it} - (k_{it-1} - k^*_{it-1}))/\nu^*$ stable, so capital is recovered from $\Delta\tilde{k}_{it}$ given $k_{i,-1} - k^*_{i,-1} = 0$. □

Equation (65) is the LQG form $\frac{1}{2}(\boldsymbol{x}_{it} - \boldsymbol{x}^*_{it})'\boldsymbol{\Theta}_i(\boldsymbol{x}_{it} - \boldsymbol{x}^*_{it})$ of the main text, with $\boldsymbol{\Theta}_i = -C^{-\gamma}Y\,\mathrm{diag}(\Theta^p_0, \tilde{\Theta}^k_0)$. Both ν^* and Θ^k_1 are negative, so $\tilde{\Theta}^k_0 > 0$.

Since $\nu^* < 0$, the target $\Delta\tilde{k}^*_{it}$ is a difference between the lagged and the current optimal capital stock. With adjustment costs the firm cares not only about the level of its capital stock but also about how costly it is to get there, and tracking $\Delta\tilde{k}^*_{it}$ captures both at once.

We verify this equivalence numerically. Drawing arbitrary stochastic sequences of actions that differ from the optimal ones, we evaluate the discounted loss in its original form, Equation (54), and in the transformed form, Equation (65), and the two agree to numerical precision at every draw.

C. Derivation of the Households' Losses from Suboptimal Actions

We approximate household j's expected discounted utility with a second-order log Taylor expansion around the non-stochastic steady state. Using the period budget constraint to substitute for consumption and expressing in log-deviations from the non-stochastic steady state yields the functional $g(\cdot)$:

$$
\begin{aligned}
&g(r_{t-1}, \pi_t, b_{jt-1}, b_{jt}, q_{jt-1}, q_{jt}, s_t, d_t, w_{jt}, w_t, l_t, t_t, \epsilon_t^c, \epsilon_t^w, \epsilon_{jt}^z) \\
= &\frac{C^{1-\gamma}\exp\{\epsilon_t^c\}}{1-\gamma}\Big(\tfrac{\omega_B}{\beta}\exp\{r_{t-1}-\pi_t+b_{jt-1}\} - \omega_B\exp\{b_{jt}\} \\
&-\omega_S\Big(\tfrac{\beta\exp\{s_t+q_{jt}\}-\beta\exp\{s_t+q_{jt-1}\}-(1-\beta)\exp\{d_t+q_{jt-1}\}}{\beta}\Big) \\
&\quad +(1+\tau_w)\omega_W\exp\{(1-\vartheta^w)w_{jt}+\vartheta^w w_t + l_t\} - \omega_T\exp\{t_t\}\Big)^{1-\gamma} \\
&-\frac{\exp\{\epsilon_t^c\}}{1-\gamma} - \frac{C^{1-\gamma}\exp\{\epsilon_t^c\}}{1+\psi}\omega_W\exp\{-\vartheta^w(1+\psi)(w_{jt}-w_t)+(1+\psi)(l_t-\epsilon_{jt}^z)\},
\end{aligned}
\tag{66}
$$

where we used $\varphi L^{1+\psi} = C^{1-\gamma}\omega_W$ and $(1-\beta)\omega_S = \beta\omega_D$ from Appendix A.3 and $\omega_B := B/C$, $\omega_T := T/C$. Let $G(\cdot)$ denote the expected discounted sum of $g(\cdot)$, obtained by multiplying the latter by β^t and summing over all t from zero to infinity, and let $\tilde{G}(\cdot)$ denote its second-order Taylor approximation around the non-stochastic steady state.

Proposition 4. *Let $\boldsymbol{x}_{jt} = (b_{jt}, w_{jt}, q_{jt})'$ denote the household's vector of actions, $\boldsymbol{z}_{jt}$ the variables taken as given, and $\boldsymbol{v}_{jt} = (\boldsymbol{x}_{jt}', \boldsymbol{z}_{jt}', 1)'$, with v_{jmt} denoting the m-th element of $\boldsymbol{v}_{jt}$. Under conditions $E_{j,-1}[b_{j,-1}^2] < \infty$, $E_{j,-1}|b_{j,-1}v_{jmt}| < \infty$ for all m, and $E_{j,-1}|v_{jmt}v_{jnt+\tau}| < \varsigma^t A$ for some $\varsigma < \beta^{-1}$ and $A \in \mathbb{R}$, the expected discounted sum of losses in utility is*

$$
\begin{aligned}
&E_{j,-1}\left[\tilde{G}(b_{j,-1}, \boldsymbol{x}_{j0}, \boldsymbol{z}_{j0}, \cdots)\right] - E_{j,-1}\left[\tilde{G}(b_{j,-1}, \boldsymbol{x}_{j0}^*, \boldsymbol{z}_{j0}, \cdots)\right] \\
&= \sum_{t=0}^{\infty}\beta^t E_{j,-1}\left[\frac{1}{2}(\boldsymbol{x}_{jt}-\boldsymbol{x}_{jt}^*)'\boldsymbol{\Theta}_0(\boldsymbol{x}_{jt}-\boldsymbol{x}_{jt}^*) + (\boldsymbol{x}_{jt}-\boldsymbol{x}_{jt}^*)'\boldsymbol{\Theta}_1(\boldsymbol{x}_{jt+1}-\boldsymbol{x}_{jt+1}^*)\right]
\end{aligned}
\tag{67}
$$

where

$$
\boldsymbol{\Theta}_0 = -C^{1-\gamma}\begin{bmatrix} \gamma\omega_B^2\left(1+\frac{1}{\beta}\right) & \gamma\omega_B\vartheta^w\omega_W & \gamma\omega_B\omega_S\left(1+\frac{1}{\beta}\right) \\ \gamma\omega_B\vartheta^w\omega_W & \vartheta^w\omega_W(1+\vartheta^w(\gamma\omega_W+\psi)) & \gamma\omega_S\vartheta^w\omega_W \\ \gamma\omega_B\omega_S\left(1+\frac{1}{\beta}\right) & \gamma\omega_S\vartheta^w\omega_W & \gamma\omega_S^2\left(1+\frac{1}{\beta}\right) \end{bmatrix}, \tag{68}
$$

$$
\boldsymbol{\Theta}_1 = C^{1-\gamma}\begin{bmatrix} \gamma\omega_B^2 & \gamma\omega_B\vartheta^w\omega_W & \gamma\omega_B\omega_S \\ 0 & 0 & 0 \\ \gamma\omega_B\omega_S & \gamma\omega_S\vartheta^w\omega_W & \gamma\omega_S^2 \end{bmatrix}. \tag{69}
$$

Proof. The proof is identical in structure to that of Proposition 1: steps "Second" to "Seventh" of Maćkowiak and Wiederholt (2015), Appendix D apply since $g(\cdot)$

depends only on current actions $\boldsymbol{x}_{jt}$, past actions $\boldsymbol{x}_{jt-1}$, and variables taken as given $\boldsymbol{z}_{jt}$. The Hessians are obtained by differentiating $g(\cdot)$ twice with respect to $\boldsymbol{x}_{jt}$, which gives $\boldsymbol{\Theta}_0$, and with respect to $(\boldsymbol{x}_{jt-1}, \boldsymbol{x}_{jt})$, which gives $\boldsymbol{\Theta}_1$ once the date-$(t+1)$ payoff is reindexed and its extra discount factor β carried along, both evaluated at the non-stochastic steady state. The share price and the dividend multiply only q_{jt} or q_{jt-1} and are zero at that steady state, so they enter the targets $\boldsymbol{x}^*_{jt}$ but not the Hessians. □

The loss Equation (67) features cross-period terms arising from asset holdings, which we remove in two steps.

Mutual fund shares. Since the mutual fund price s_t adjusts to clear the share market, it is always optimal for households to hold a constant share in deviations from steady state: $q^*_{jt} = 0$. Moreover, the loss depends on bonds and shares only through the combination $\omega_B b_{jt} + \omega_S q_{jt}$. The share choice is therefore trivial and q_{jt} can be dropped from the action vector, reducing it to $(b_{jt}, w_{jt})'$ with Hessians given by the upper-left 2×2 blocks of $\boldsymbol{\Theta}_0$ and $\boldsymbol{\Theta}_1$.

Bond holdings. Following Maćkowiak and Wiederholt (2015), bond holdings b_{jt} can be substituted out using the budget constraint. Deviations of bond holdings from their perfect-information counterpart satisfy

$$\omega_B(b_{jt} - b^*_{jt}) = \tfrac{1}{\beta}\omega_B(b_{jt-1} - b^*_{jt-1}) - \left[(c_{jt} - c^*_{jt}) + \vartheta^w \omega_W (w_{jt} - w^*_{jt})\right]. \tag{70}$$

This recursion, together with $b^*_{j,-1} = b_{j,-1}$ and the transversality condition, allows bond deviations to be expressed entirely in terms of current and past consumption and wage deviations. The household's action vector then becomes $(c_{jt}, w_{jt})'$, as in the main text: with bonds substituted out, choosing how much to save is the same as choosing how much to consume.

Proposition 5. *Under the same conditions, Equation (67) equals*

$$-C^{1-\gamma} \sum_{t=0}^{\infty} \beta^t E_{j,-1}\left[\frac{\gamma}{2}(c_{jt} - c^*_{jt})^2 + \frac{\vartheta^w \omega_W (1 + \vartheta^w \psi)}{2}(w_{jt} - w^*_{jt})^2\right], \tag{71}$$

giving the Hessian

$$\boldsymbol{\Theta}_j = -C^{1-\gamma} \begin{bmatrix} \gamma & 0 \\ 0 & \vartheta^w \omega_W (1 + \vartheta^w \psi) \end{bmatrix}. \tag{72}$$

Proof. Define $\Delta_t := \omega_B(b_{jt} - b^*_{jt})$, and decompose $\Delta_t = \Delta_{c,t} + \Delta_{w,t}$ where $\Delta_{c,t}$ and $\Delta_{w,t}$ accumulate past consumption and wage deviations respectively, with $\Delta_{c,-1} = \Delta_{w,-1} = 0$. Substituting into Equation (67), multiplying by β^t, summing from $t = 0$ to T, taking $E_{j,-1}$, and passing to the limit $T \to \infty$ — using the stated conditions to eliminate boundary terms and $b^*_{j,-1} = b_{j,-1}$ — yields Equation (71) by telescoping. See Maćkowiak and Wiederholt (2015), Appendix D, for the complete argument. □

The optimal actions under perfect information are given in Appendix A.4.

We verify this equivalence numerically, as we did for firms: the discounted loss in Equation (67) and in Equation (71) agree to numerical precision for arbitrary sequences of actions.

D. Computing the Equilibrium

We solve for the equilibrium in the sequence space, iterating on impulse response functions over T periods. We assume the following initial conditions: $k_{i,-1} = 0$ for every firm $i \in [0, 1]$, $b_{j,-1} = q_{j,-1} = 0$ for every household $j \in [0, 1]$, and initial beliefs at the steady state of the Kalman filter for every decision-maker. Because the dynamic responses of p_t, k_t, c_t and w_t, together with those of the exogenous processes included in the solve, determine every other aggregate variable, we only iterate on $\boldsymbol{\mathcal{J}} = (\mathcal{J}_p, \mathcal{J}_k, \mathcal{J}_c, \mathcal{J}_w)$.[18] Recovering the full moving-average representation requires solving for all seven shocks whereas the estimation procedure of Section 5 requires only the monetary policy process. The fixed point requires consistency between the perceived and the actual law of motion which, here, depends on the signals chosen by firms and households. When we represent that law of motion as a VARMA, we recover it from the impulse responses using the frequency-domain projection method of Han, Tan, and Wu (2022); when we represent it as a truncated moving average, it follows directly from the impulse responses themselves. We solve the attention problems with the algorithm of Miao, Wu, and Young (2022).

1. Guess $\boldsymbol{\mathcal{J}}^{(n)}$.
2. Compute the firms' targets p^*_{it} and k^*_{it} implied by $\boldsymbol{\mathcal{J}}^{(n)}$, approximate the joint process $(\Delta\zeta^*_t, k^*_t)$ with a finite-order VARMA, and build the state-space representation.
3. Solve the firms' attention problem and compute the implied $\mathcal{J}_p$ and $\mathcal{J}_k$.
4. Compute the households' targets c^*_{jt} and w^*_{jt} given the updated real rate and labor, and build the corresponding state-space representation.

[18] One could swap the price level and capital responses for those of inflation and investment, but this typically requires a tighter convergence tolerance.

5. Solve the households' attention problem and compute the implied $\mathcal{J}_c$ and $\mathcal{J}_w$.
6. Collect the implied responses in $\boldsymbol{\mathcal{J}}^{(n)}_{\text{imp}}$. If $\|\boldsymbol{\mathcal{J}}^{(n)}_{\text{imp}} - \boldsymbol{\mathcal{J}}^{(n)}\|_\infty > \tau$, set $\boldsymbol{\mathcal{J}}^{(n+1)}$ by linear interpolation between the two and return to step 2.

Implementation. The fixed point is computed on an impulse-response horizon of $T = 400$ quarters. Within each iteration, the equilibrium target processes are realized as moving averages with an effective horizon of 60 quarters, the attention problems are solved on that representation, and the resulting action paths are smoothly merged into their targets at the tail.[19] Iterates are updated with an interpolation weight of 0.10 on the new solution. For each block — prices, capital, consumption, and the real wage — the residual is the sum of squared revisions to its path over the first 41 horizons, with paths measured as responses to a one-standard-deviation shock and scaled by 100. The algorithm stops when the largest of the four residuals falls below the tolerance $\tau = 10^{-4}$, so the residuals at every reported fixed point are bounded by these values by construction.

Numerical accuracy. The settings above deliver accurate solutions. Tightening the tolerance to 10^{-6} changes the impulse responses by less than 0.08 percentage points, extending the convergence window to 200 quarters leaves them unchanged, and the same fixed point is reached for any interpolation weight between 0.005 and 0.10.

During the MCMC estimation procedure, a share of proposed draws — roughly a third — fails to converge. These are not failures of the solver: they are parameter combinations that are not economically admissible, implying firms or households that pay essentially no attention, at which point the fixed point has no settled solution. They are analogous to a habit parameter above one in the NK-DSGE — a region the sampler must be kept out of — except that the admissible region cannot be characterized a priori and is instead enforced by the convergence check itself. Such draws arise because the impulse responses on which the model is estimated require a great deal of inertia, pushing proposals toward the low-attention boundary; non-convergent proposals are rejected, which slows mixing but does not bias the retained sample.

[19] The merge leaves the rational-inattention solution itself unaffected: it operates at horizons far enough that actions and optimal actions coincide, and only prevents the actions from inheriting the mechanical reversion to zero of the truncated moving-average realization.

E. NK-DSGE: Linearized Equilibrium Conditions.

We log-linearize the optimality conditions around the non-stochastic steady state and obtain the following system of thirteen equations:

$$\Lambda_t = E_t[\Lambda_{t+1}] - \tfrac{1}{\gamma} E_t[r_t - \pi_{t+1}] \tag{73}$$

$$\Lambda_t = \tfrac{\gamma h c_{t-1}}{(1-h)(1-\beta h)} - \tfrac{\gamma(1+\beta h^2)c_t}{(1-h)(1-\beta h)} + \tfrac{\gamma \beta h E_t[c_{t+1}]}{(1-h)(1-\beta h)} + \tfrac{\gamma(1-\beta h \rho_c)\epsilon_t^c}{(1-h)(1-\beta h)} \tag{74}$$

$$\begin{aligned} k_t = {} & \tfrac{k_{t-1}}{(1+\beta)} + \tfrac{\beta E_t[k_{t+1}]}{(1+\beta)} + \tfrac{[1-\beta(1-\delta)]E_t[w_{t+1}]}{\chi_i(1+\beta)} + \tfrac{[1-\beta(1-\delta)]E_t[mpk_{t+1}]}{\chi_i(1+\beta)} \\ & - \tfrac{[1-\beta(1-\delta)]E_t[mpl_{t+1}]}{\chi_i(1+\beta)} - \tfrac{[1-\beta(1-\delta)\rho_i]\epsilon_t^i}{\chi_i(1+\beta)} + \tfrac{E_t[\Lambda_{t+1}]}{\chi_i(1+\beta)} - \tfrac{\Lambda_t}{\chi_i(1+\beta)} \end{aligned} \tag{75}$$

$$\pi_t = \kappa_p mc_t + \beta E_t[\pi_{t+1}] + \hat{\epsilon}_t^p \tag{76}$$

$$w_t = \kappa_w \widehat{mrs}_t + \tfrac{w_{t-1}}{(1+\beta)} - \tfrac{\pi_t}{(1+\beta)} + \tfrac{\beta E_t[w_{t+1}]}{(1+\beta)} + \tfrac{\beta E_t[\pi_{t+1}]}{(1+\beta)} + \hat{\epsilon}_t^w \tag{77}$$

$$mpl_t = a_t + (\alpha - 1) l_t + \phi k_{t-1} \tag{78}$$

$$mpk_t = a_t + \alpha l_t + (\phi - 1) k_{t-1} \tag{79}$$

$$mc_t = w_t - mpl_t \tag{80}$$

$$\widehat{mrs}_t = \gamma c_t + \psi l_t - w_t \tag{81}$$

$$i_t = \tfrac{1}{\delta} k_t - \tfrac{1-\delta}{\delta} k_{t-1} - \epsilon_t^i \tag{82}$$

$$y_t = \tfrac{C}{Y} c_t + \tfrac{I}{Y} i_t + \tfrac{G}{Y} \epsilon_t^g \tag{83}$$

$$y_t = a_t + \alpha l_t + \phi k_{t-1} \tag{84}$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\phi_\pi \pi_t + (1-\rho_r)\phi_{y^*} y_t^* + \epsilon_t^m \tag{85}$$

where κ_p and κ_w are the slopes of the price and wage Phillips curves respectively, and Λ_t is the marginal utility of consumption. We solve the system of linear rational expectations comprising Equation (73)–Equation (85) together with the exogenous stochastic processes using Sims (2002) `gensys`.

E.1. Firms' Equilibrium Conditions to an Idiosyncratic Shock. The following conditions determine the individual actions of a firm i subject to Rotemberg price adjustment costs and investment adjustment costs, in response to an idiosyncratic productivity shock ϵ_{it}^a; aggregate variables are unaffected.

$$
\begin{aligned}
p_{it} &= (1+\kappa_p+\beta)^{-1}\left[p_{it-1}+\kappa_p mc_{it}+\beta E_{it}[p_{it+1}]\right] && (86)\\
mc_{it} &= (1-\alpha)l_{it}-\phi k_{it-1}-\epsilon_{it}^a && (87)\\
v_{it} &= (1-\beta(1-\delta))\,E_{it}[mpk_{it+1}+mc_{it+1}]+\beta(1-\delta)E_{it}[v_{it+1}] && (88)\\
mpk_{it} &= \epsilon_{it}^a+(\phi-1)k_{it-1}+\alpha l_{it} && (89)\\
-\vartheta^p p_{it} &= \epsilon_{it}^a+\phi k_{it-1}+\alpha l_{it} && (90)\\
i_{it} &= \tfrac{v_{it}}{(1+\beta)\chi_i}+\tfrac{i_{it-1}}{(1+\beta)}+\tfrac{\beta E_{it}[i_{it+1}]}{(1+\beta)} && (91)\\
i_{it} &= \tfrac{1}{\delta}k_{it}-\tfrac{1-\delta}{\delta}k_{it-1} && (92)
\end{aligned}
$$

where the Rotemberg price adjustment cost parameter χ_p can be inferred from the slope of the New Keynesian Phillips curve, $\kappa_p=\frac{\vartheta^p-1}{\chi_p}$, and v_{it} is the relative marginal value of firm i's installed capital.

F. Data

Table 4 details the construction of each series used in the estimations.

G. Empirical Responses to a Monetary Policy Shock

In this appendix we describe how we estimate the empirical responses that serve as the target of the direct estimation: we project each observable on the monetary policy shock of Romer and Romer (2004), following Jordà (2005). We detail the projection first and then the construction of the weighting matrix.

G.1. Estimation. The projection is run at monthly frequency: the four matched variables — consumption, investment, the real wage, and the price level — are temporally disaggregated, using the monthly indicators listed in Table 4.[20] The shock series is divided by four to express it in quarterly rate units and standardized before the projection, so that $\widehat{\mathcal{J}}_h$ is the response to a one-standard-deviation shock.

For each observable y and horizon $h=1,\dots,H$ with $H=48$ months, we estimate by OLS the projection

$$
y_{t+h}-y_{t-1}=\mathcal{J}_h\,\varepsilon_t^m+\boldsymbol{\beta}_h'\mathbf{X}_t+\epsilon_{t,h}, \tag{93}
$$

[20]The disaggregation is Chow–Lin. Consumption and investment are disaggregated component by component and summed afterwards, and the indicators for real quantities are deflated by PCEPI beforehand.

so that $\widehat{\mathcal{J}}_h$ is read directly as the impulse response at horizon h. The control vector $\mathbf{X}_t$ follows Ramey (2016): the contemporaneous values of industrial production, the unemployment rate, the consumer price index and the commodity price index (IP, U, CPI and PCOM in Table 4); two monthly lags of those four series and of the federal funds rate; and two lags of the shock. Standard errors are computed with the Newey and West (1987) correction, with bandwidth $h + 1$ at horizon h. The monthly responses are aggregated to sixteen quarterly horizons by averaging within each quarter, without further smoothing.

G.2. Weighting Matrix. $\Sigma_{\widehat{\mathcal{J}}}$ is the joint sampling covariance of the four responses across the sixteen horizons, estimated by a wild residual bootstrap Montiel Olea and Plagborg-Møller (2021) and regularized by shrinkage toward its diagonal. The estimated shrinkage parameter is negligible, so the weighting matrix is essentially the raw bootstrap covariance.

H. Quasi-Bayesian Priors

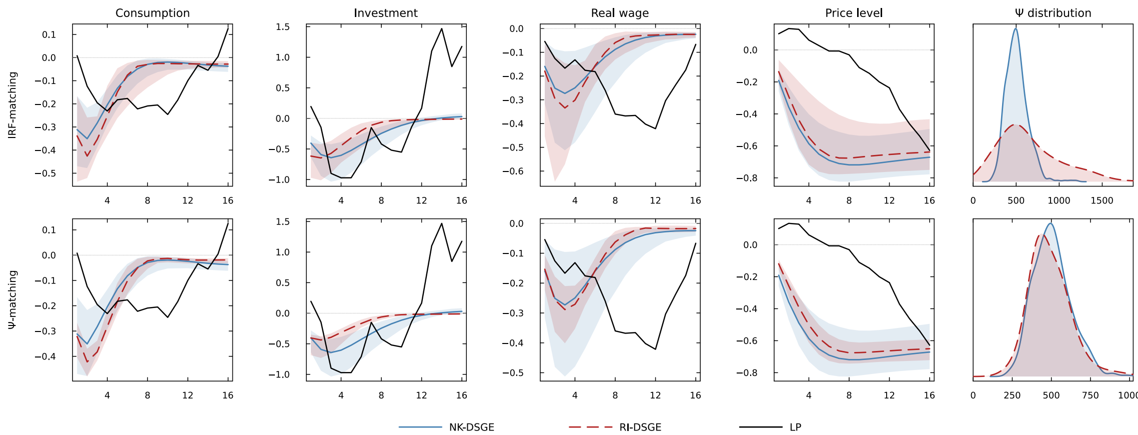
This appendix details the priors of the quasi-Bayesian estimations. The NK-DSGE parameters are assigned the priors of Table 6. The three RI parameters are positive, so we assign log-normal priors to all three.

For the direct estimation we select the six hyperparameters — the median and the log-standard deviation of each of λ_i, λ_j and χ_k — so that the RI prior says the same thing as the NK prior, under the two criteria reported in Table 5.

The first matches dynamics. Drawing from the NK prior and solving the model, we record the fifth, fiftieth and ninety-fifth percentile paths of the four responses at every horizon. We then pick the RI hyperparameters whose implied response distribution reproduces those three paths. The criterion never touches the data: the two priors are made to say the same thing about dynamics, and any posterior divergence is attributable to the likelihood alone. Figure 12 reports the two prior-predictive distributions. The two sets of bands are close at every horizon and for all four variables: matching the three percentile paths disciplines the location, width and decay of the prior-predictive responses. The construction nonetheless leaves the implied distribution of fit more dispersed than the NK model's — the prior-predictive Ψ has a 90% range of $[189, 1371]$ against the NK model's $[335, 734]$ — and in particular a fatter left tail of good fits (fifth percentile of 189 against 335), the head start the second criterion is designed to remove.

The second matches fit. Evaluating Equation (21) at each draw yields a prior-predictive distribution of the criterion $\Psi(\boldsymbol{\theta})$, and we pick the RI hyperparameters

Figure 12: Prior-predictive distributions under the two RI prior constructions



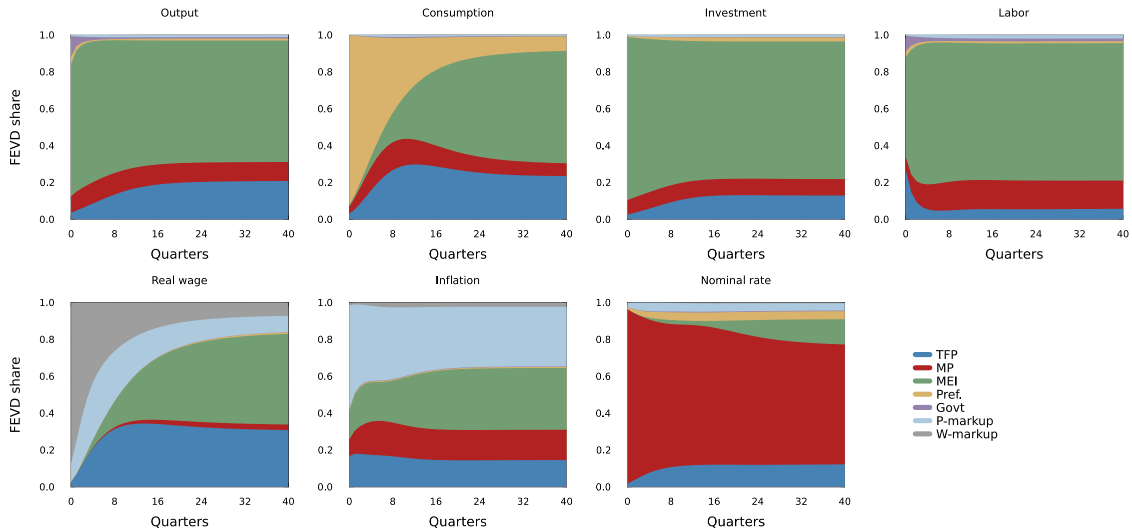
that reproduce the NK model's quantile by quantile. This is the more conservative construction: it equalizes the two models' a priori capacity to fit the target, so that the RI prior cannot carry a head start — a fatter left tail of good fits — of the kind the first criterion turns out to allow. The match is tight: the prior-predictive Ψ has median 453 and 90% range $[337, 707]$ under the fitted RI prior, against 498 and $[335, 734]$ for the NK model. Its cost is that it depends on the target and must be refit whenever the target changes. Because it is calibrated against $\widehat{\mathcal{J}}$, this prior is data-dependent and the marginal data densities it delivers are not ex ante Bayes factors. We therefore report the headline comparison under the IRF-matched prior, which never touches the data, and read the Ψ-matched one as a sensitivity check.

Both sets of priors are computed the same way. We draw 1,000 combinations of RI parameters from a wide proposal and solve the model at each of them. We then search for the hyperparameters that match each criterion, reweighting the solved draws by importance.

The indirect estimation requires no such discipline. Models are compared by their credible bands and likelihood, not by marginal data densities, so the prior plays no evidential role there. We accordingly assign flat priors over the log parameter space and let the criterion alone shape the quasi-posterior.[21]

[21]The prior still matters through the region it lets the sampler reach. A flat prior is what we want here: the RI-DSGE's capacity to reproduce the NK responses should be limited by the model, not by a prior we chose ourselves.

Figure 13: Forecast Error Variance Decomposition



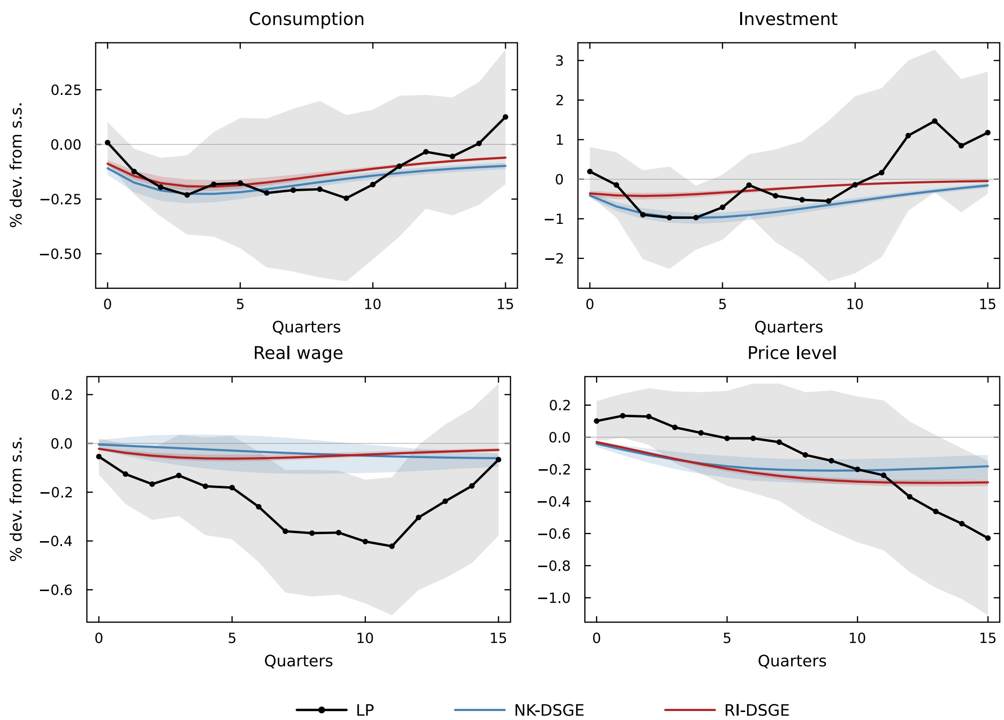
I. Full-Information Bayesian Estimation

We use Bayesian methods to characterize the posterior distribution of the NK-DSGE frictions and of the parameters of the exogenous stochastic processes on the full covariance of aggregate time series. The likelihood is computed with the Kalman filter applied to the seven observables $\boldsymbol{\mathcal{Y}}_t = [Y_t,\ C_t,\ I_t,\ H_t,\ W_t,\ \pi_t,\ R_t]'$, whose construction is described in Appendix F. The observables are quarterly and span 1966Q1–2018Q4; monthly series are aggregated to the quarter as within-quarter means for the federal funds rate and as end-of-quarter values for population and employment. All series are detrended with the HP filter, smoothing parameter 1600, within the sample, and we retain the cyclical component.

The chain is run for 120,000 draws with an acceptance rate of 0.289. Table 6 reports the priors and the resulting posterior.

J. Quasi-Bayesian Posteriors

Figure 14: Empirical impulse responses to a monetary policy shock vs. model fit



Note: Black dotted lines: local-projection responses to a one-standard-deviation Romer and Romer (2004) shock, with 90% confidence bands in grey. Colored lines and bands: quasi-posterior median responses and 90% credible bands of the RI-DSGE (red) and NK-DSGE (blue), under the Ψ-matched prior.

Figure 15: RI-DSGE quasi-posterior parameter distributions (IRF-matched prior)

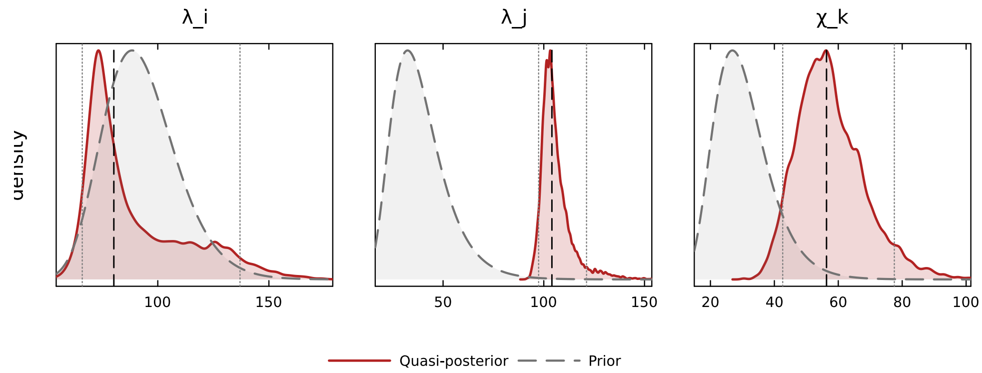


Figure 16: RI-DSGE quasi-posterior parameter distributions (Ψ-matched prior)

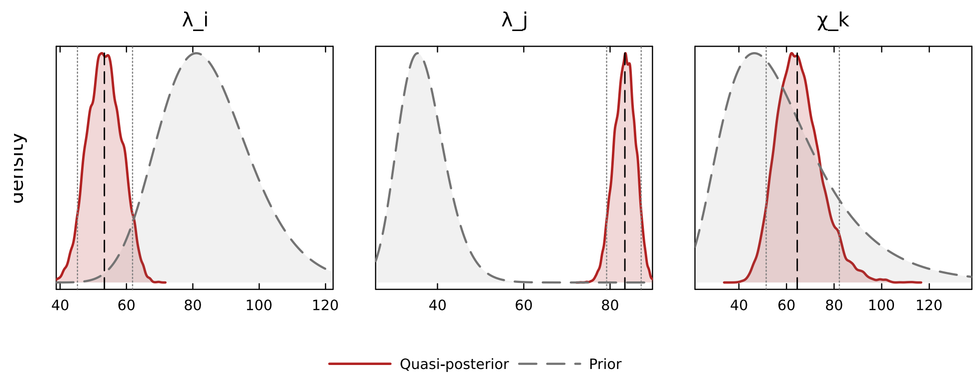


Figure 17: RI-DSGE quasi-posterior parameter distributions (indirect estimation)

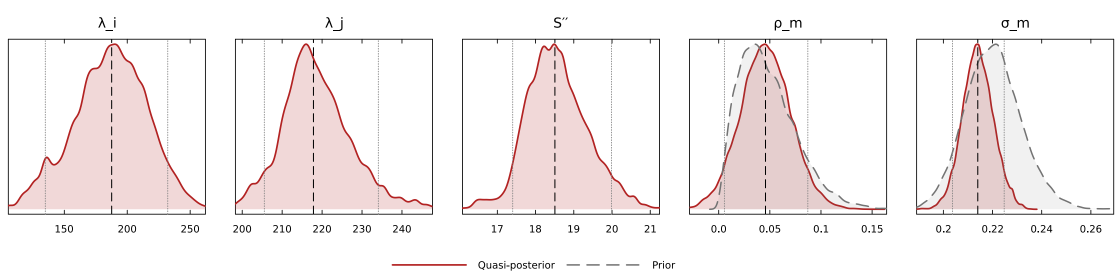


Figure 18: RI-DSGE responses to all shocks

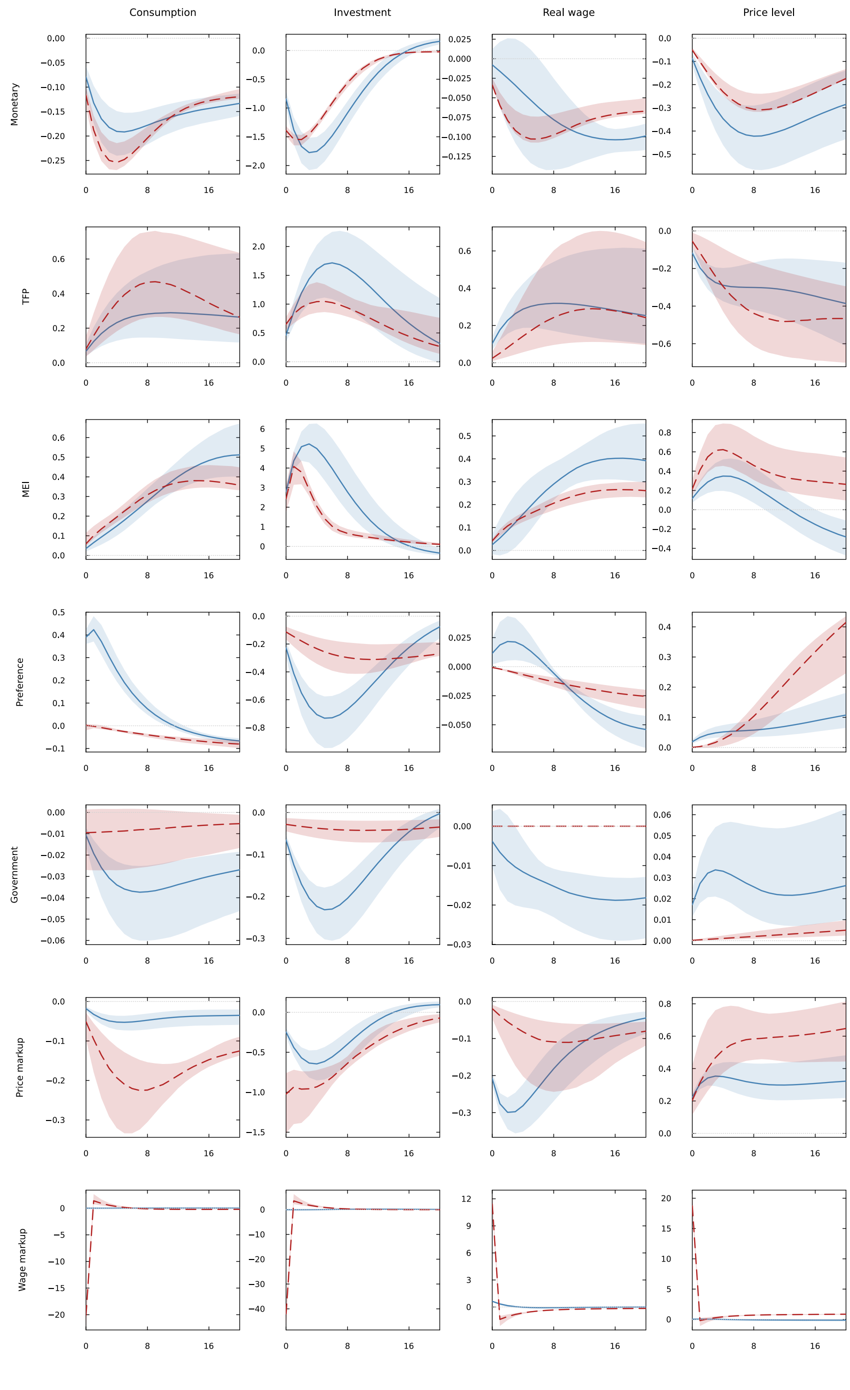


Table 4: Description of the Data

Panel A: Series construction

Consumption	$C = (\text{CND} + \text{CS}) \times Y$
Investment	$I = (\text{CD} + \text{GPDI}) \times Y$
Real Wage	$W = \text{HC}$
Price Level	$P = \text{DEF}$
Output	$Y = \text{GDP}$
Hours Worked	$H = (\text{WH} \times \text{EMP})/\text{POP}$
Inflation Rate	$\pi = \log(\text{P}) - \log(\text{P}_{-1})$
Nominal Interest Rate	$R = \text{FFR}/4$

Panel B: Raw series

Symbol	Mnemonic	Description
Quarterly		
GDP	A939RX0Q048SBEA	Real GDP per capita
CND	DNDGRE1Q156NBEA	Nondurables, GDP share
CS	DSERRE1Q156NBEA	Services, GDP share
CD	DDURRE1Q156NBEA	Durables, GDP share
GPDI	A006RE1Q156NBEA	Private investment, GDP share
HC	COMPRNFB	Real hourly compensation
DEF	GDPDEF	GDP deflator
WH	PRS85006023	Average weekly hours
Monthly		
EMP	CE16OV	Employment level
POP	CNP16OV	Population level
FFR	FEDFUNDS	Federal funds rate
IP	INDPRO	Industrial production
U	UNRATE	Unemployment rate
CPI	CPIAUCSL	Consumer price index
PCOM	PPICRM	PPI, crude materials
ε^m	—	Monetary policy shock
Monthly indicators		
CND^m	PCEND	PCE, nondurables
CS^m	PCES	PCE, services
CD^m	PCEDG	PCE, durables
I^m	IPBUSEQ	IP, business equipment
W^m	AHETPI	Average hourly earnings
P^m	PCEPI	PCE price index

Notes: Mnemonics refer to the FRED database (fred.stlouisfed.org); ε^m is the original residual series from the data appendix of Romer and Romer (2004). Quantities are per capita and all matched series are in $100 \cdot \log$. The first four variables in Panel A are those matched in the estimation; the remaining ones enter only the full-information estimation of the NK-DSGE. IP, U, CPI and PCOM are the local-projection controls of Equation (93). The monthly indicators are used to disaggregate their corresponding quarterly series; the real-quantity indicators are deflated by PCEPI before use.

Table 5: Prior Distributions, RI-DSGE

Parameter		Density	Mean	Std.	[5th, 95th]
IRF-matched					
λ_i	Firms' marginal cost of attention	LN	92.8	16.7	[68.0, 122.6]
λ_j	Households' marginal cost of attention	LN	38.8	13.8	[20.7, 64.5]
χ_k	Capital adjustment cost	LN	30.4	9.0	[18.2, 46.9]
Ψ-matched					
λ_i	Firms' marginal cost of attention	LN	84.4	14.0	[63.5, 109.3]
λ_j	Households' marginal cost of attention	LN	36.6	5.4	[28.4, 46.2]
χ_k	Capital adjustment cost	LN	59.0	24.5	[28.3, 105.1]

Note: LN = Log-Normal. Means, standard deviations and percentiles in levels. The IRF-matched priors are centered so that the prior-predictive impulse responses resemble those implied by the NK priors; the Ψ-matched priors so that the prior-predictive distribution of the criterion does. The NK-DSGE priors are those of Table 6.

Table 6: Prior and Posterior Distributions

		Priors			Posteriors			
Parameter		Density	Mean	Std.	Mode	Median	Std.	[5th, 95th]
	New Keynesian frictions							
h	Consumption habit	N	0.5	0.1	0.87	0.87	0.02	[0.84, 0.90]
χ_i	Investment adjustment costs	G	4.0	1.0	3.83	3.94	0.58	[3.11, 5.01]
κ_p	NKPC slope — prices	G	0.1	0.05	0.05	0.05	0.01	[0.03, 0.08]
κ_w	NKPC slope — wages	G	0.1	0.05	0.01	0.01	0.00	[0.00, 0.01]
	Exogenous processes							
ρ_a	TFP	B	0.6	0.2	0.93	0.92	0.05	[0.83, 0.98]
ρ_m	M.P.	B	0.4	0.2	0.03	0.04	0.03	[0.01, 0.10]
ρ_i	MEI	B	0.6	0.2	0.77	0.75	0.06	[0.65, 0.83]
ρ_c	Risk Prem.	B	0.6	0.2	0.27	0.27	0.07	[0.17, 0.39]
ρ_g	Govt	B	0.6	0.2	0.81	0.81	0.04	[0.74, 0.87]
ρ_p	Price markup	B	0.6	0.2	0.76	0.74	0.06	[0.64, 0.83]
ρ_w	Wage markup	B	0.6	0.2	0.73	0.72	0.05	[0.62, 0.79]
θ_p	Price markup MA	B	0.5	0.2	0.37	0.34	0.11	[0.15, 0.52]
θ_w	Wage markup MA	B	0.5	0.2	0.81	0.81	0.03	[0.77, 0.86]
$100\sigma_a$	TFP	I	0.5	1.0	0.63	0.63	0.03	[0.58, 0.68]
$100\sigma_m$	M.P.	I	0.1	1.0	0.22	0.22	0.01	[0.20, 0.24]
$100\sigma_i$	MEI	I	0.5	1.0	4.82	5.21	0.78	[4.14, 6.71]
$100\sigma_c$	Risk Prem.	I	0.1	1.0	3.16	3.25	0.45	[2.69, 4.17]
$100\sigma_g$	Govt	I	0.5	1.0	1.47	1.48	0.07	[1.37, 1.60]
$100\sigma_p$	Price markup	I	0.1	1.0	0.15	0.16	0.02	[0.13, 0.19]
$100\sigma_w$	Wage markup	I	0.1	1.0	0.48	0.49	0.03	[0.45, 0.54]

Notes: N = Normal, B = Beta, G = Gamma, I = Inverted-Gamma. [5th, 95th] are posterior percentiles. The price- and wage-markup shocks are rescaled to enter their Phillips curves with a coefficient of one.

Figure 19: RINK responses to a monetary policy shock

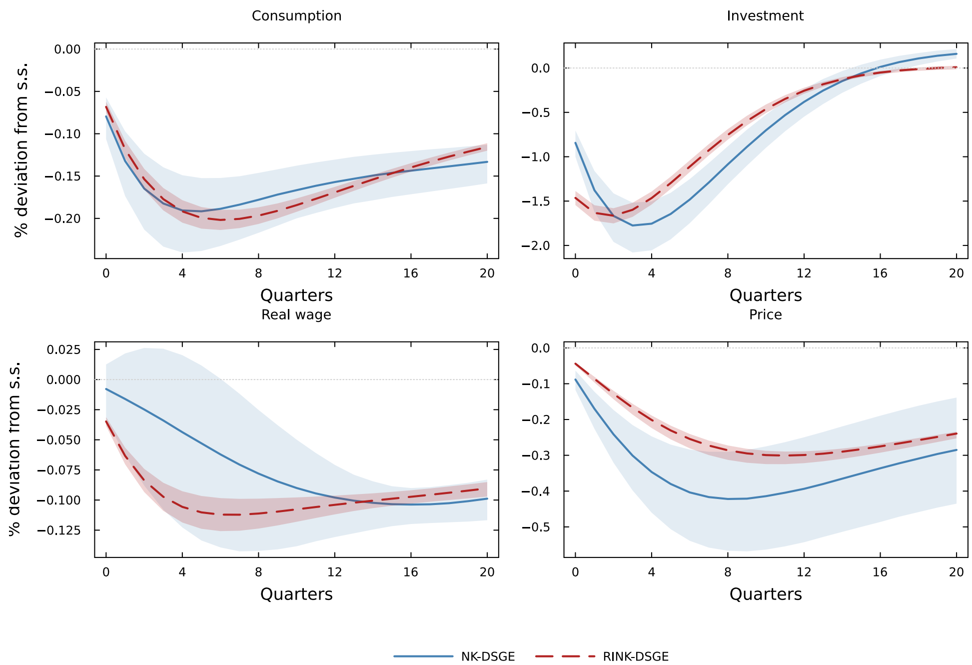


Figure 20: RINK quasi-posterior parameter distributions

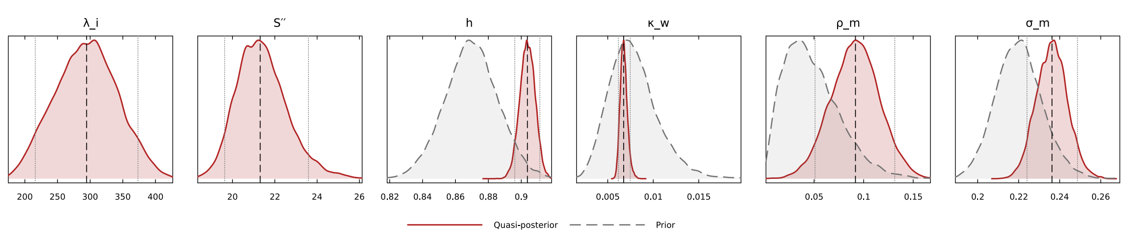


Figure 21: RINK responses to all shocks

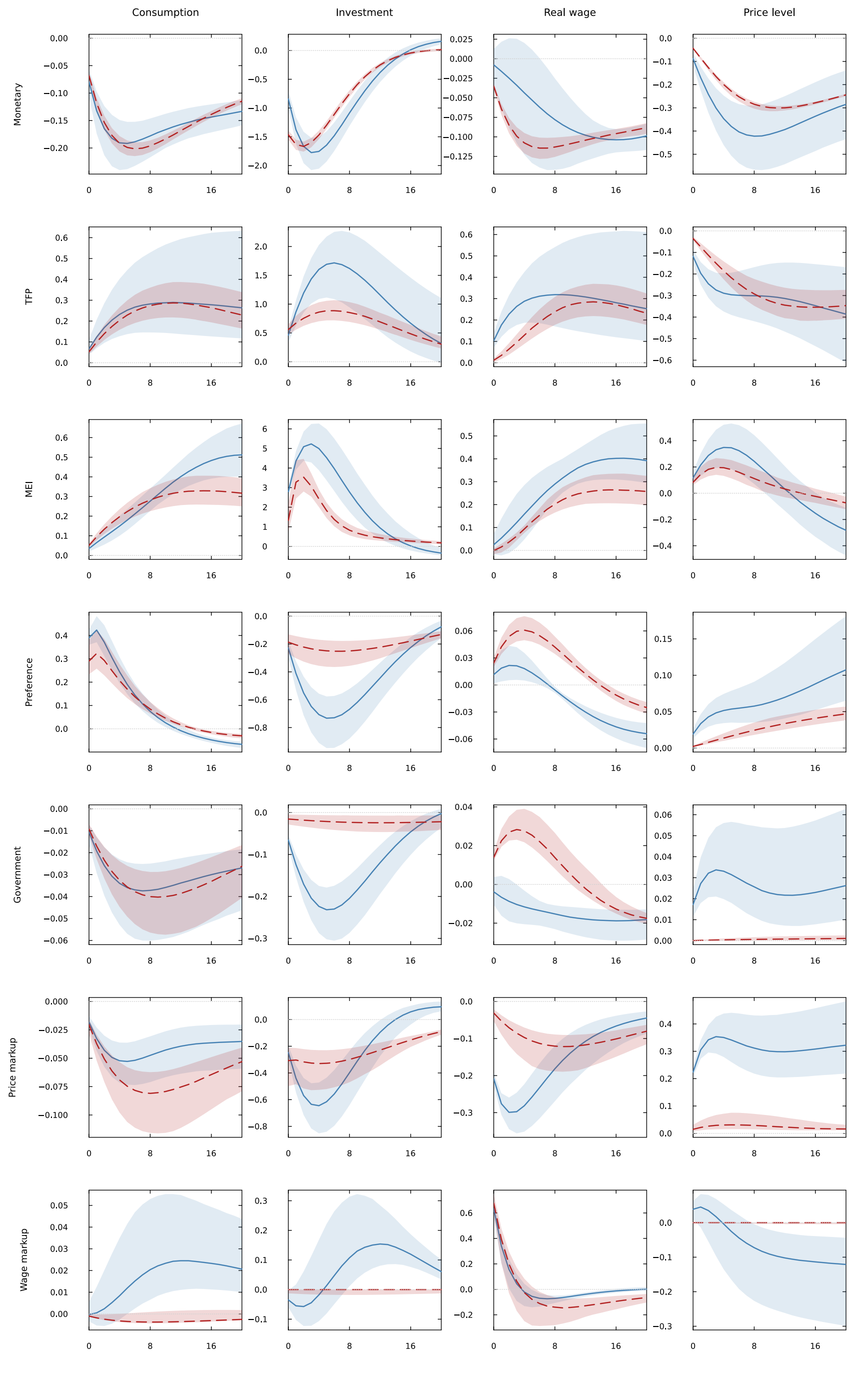


Figure 22: NKRI responses to a monetary policy shock

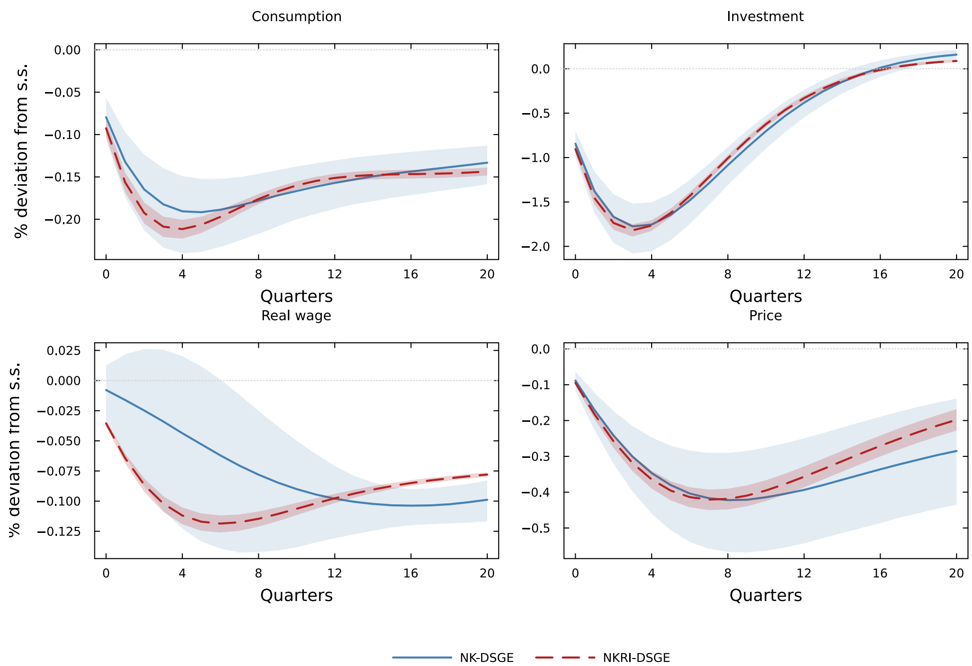


Figure 23: NKRI quasi-posterior parameter distributions

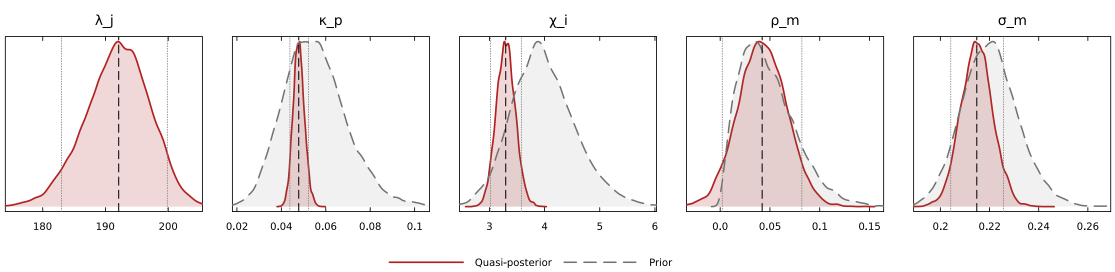


Figure 24: NKRI responses to all shocks

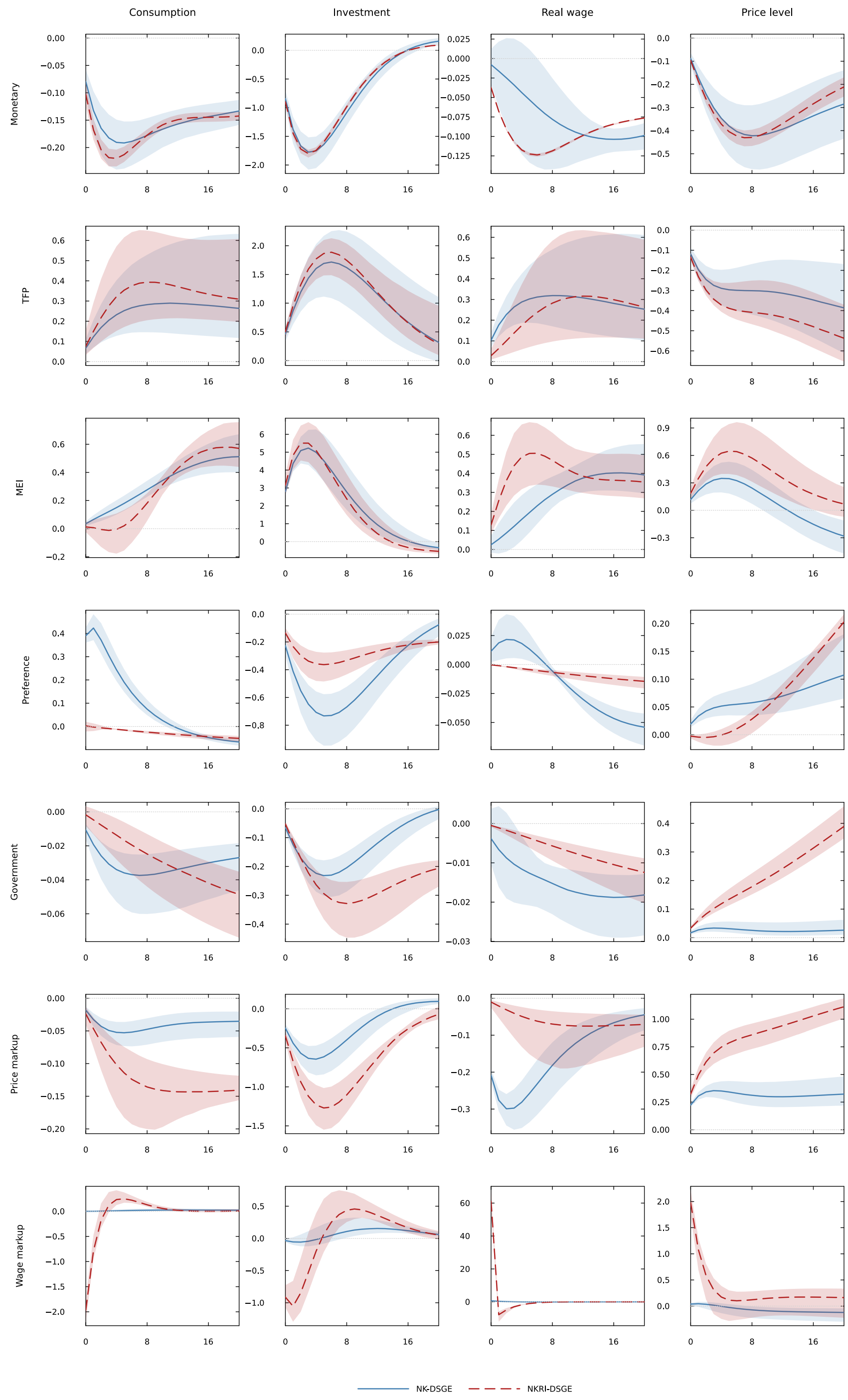


Table 7: Quasi-Posterior Summaries and Convergence Diagnostics

Parameter	Median	[5^{th}, 95^{th}]	$\widehat{R}$	ESS
Panel A: Direct Estimation				
RI-DSGE, IRF-matched prior				
λ_i	80.1	[65.9, 137.2]	1.04	1216
λ_j	104.1	[97.5, 120.4]	1.03	1566
χ_k	56.2	[42.5, 77.7]	1.02	2117
RI-DSGE, Ψ-matched prior				
λ_i	53.3	[45.2, 61.8]	1.01	3248
λ_j	83.4	[79.2, 87.2]	1.03	4050
χ_k	64.5	[51.5, 82.2]	1.02	4145
Panel B: Indirect Estimation				
RI-DSGE				
λ_i	187.63	[134.89, 232.21]	1.015	1083
λ_j	217.84	[205.53, 234.08]	1.042	275
χ_k	18.51	[17.40, 19.99]	1.010	873
ρ_m	0.046	[0.006, 0.087]	1.005	1483
σ_m	0.214	[0.204, 0.225]	1.008	1233
NKRI-DSGE				
λ_j	192.11	[183.01, 199.86]	1.001	2,426
κ_p	0.048	[0.044, 0.052]	1.001	3,872
χ_i	3.30	[3.02, 3.58]	1.001	3,845
ρ_m	0.042	[0.002, 0.082]	1.001	3,972
σ_m	0.215	[0.204, 0.226]	1.001	3,445
RINK-DSGE				
λ_i	294.58	[215.92, 373.23]	1.001	3,451
χ_k	21.31	[19.63, 23.59]	1.004	3,071
h	0.904	[0.896, 0.911]	1.002	3,267
κ_w	0.0067	[0.0062, 0.0075]	1.001	2,489
ρ_m	0.092	[0.051, 0.131]	1.001	3,141
σ_m	0.236	[0.224, 0.249]	1.002	3,168

Notes: Each quasi-posterior pools at least eight chains of 8,000 retained draws started from over-dispersed points. $\widehat{R}$ is computed across the eight chains; ESS is the Geyer initial-positive-sequence effective sample size of the pooled draws.