Macroeconomic Dynamics (2024), **28**, pp. 1371–1393
doi:10.1017/S1365100523000500

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Linear quadratic approximation of rationally inattentive control problems†

Jianjun Miao[1] and Bo Zhang[2]

[1]Department of Economics, Boston University, Boston, MA, USA
[2]Institute of Chinese Financial Studies, Southwestern University of Finance and Economics, Chengdu, China
Corresponding author: Jianjun Miao; Email: miaoj@bu.edu

Abstract

This paper proposes a linear quadratic approximation approach to dynamic nonlinear rationally inattentive control problems with multiple states and multiple controls. An efficient toolbox to implement this approach is provided. Applying this toolbox to five economic examples demonstrates that rational inattention can help explain the comovement puzzle in the macroeconomics literature.

Keywords: Rational inattention; information theory; tracking problem; optimal control; entropy; numerical methods

1. Introduction

People do not pay attention to all available information because processing information is costly. Sims (1998, 2003, 2011) introduces rational inattention (RI) models that study how people should optimize when their abilities to translate external data into action are constrained by a finite Shannon capacity to process information. Such models can generate sluggish and inertia responses to external information without introducing frictions like adjustment costs and thus have wide applications in macroeconomics.[1]

Despite the rapid growth of this literature, a major hurdle for beginners is the difficulty of solving general multivariate nonlinear control problems under RI. Sims (2003, 2011) formulates such problems in the linear-quadratic Gaussian (LQG) framework. Using dynamic semidefinite programming in this framework, Miao, Wu, and Young (henceforth MWY) (2022) provide a characterization of the optimal solution under RI and both the value function-based and first-order conditions-based numerical methods to compute such a solution. However, little is known about how to solve general multivariate nonlinear rationally inattentive control problems.

The goal of this paper is to fill this gap in the literature by making three contributions. First, we develop an LQ approximation approach that consists of two broad steps. In step 1, we take the decision-maker (DM)'s information structure as exogenously given and then solve the nonlinear control problem under partial information by LQ approximation. The approximated linear policy function is the same as that under full information by the certainty equivalence principle. In step 2, we use the methods of MWY (2022) to solve for the optimal information structure. Under Gaussian uncertainty, the optimal information structure can be represented by a signal vector that is a linear transformation of hidden states plus a noise. Both the linear transformation and the noise covariance matrix are endogenously chosen.

†We are grateful to two anonymous referees and an associate editor for helpful comments. Bo Zhang is grateful to the support by the Guanghua Talent Project of Southwestern University of Finance and Economics.

© The Author(s), 2023. Published by Cambridge University Press.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

Second, we develop a Matlab toolbox to efficiently compute LQ approximations for multivariate nonlinear rationally inattentive control problems. This toolbox is robust and easy to apply in practice. It delivers the same approximated linear policy function based on the perturbation approach as in the dynamic stochastic general equilibrium (DSGE) literature computed by a number of public packages, for example, Dynare.

Third, we apply our toolbox to five economic examples that can be formulated as rationally inattentive nonlinear control problems with multivariate states and multivariate controls. In a PC with Intel Core i7-10700 CPU and 64 GB memory, it takes less than 1 s to solve each of these examples for a wide range of parameter values. The first example illustrates the pitfall of the ad hoc LQ approximation method. The next three examples consider a rationally inattentive social planner's resource allocation problems with various sources of exogenous shocks. These decision problems come from the real business cycles (RBC) literature, though we do not study decentralized market equilibrium. We use these examples to illustrate how RI combined with variable capital utilization can generate comovement among investment, consumption, and labor hours in response to various shocks to the demand side of the economy. In the last example, we study a consumption/saving problem with both durable and nondurable goods similar to Luo et al. (2015). The utility function takes a general power form beyond a quadratic function. This example shows that RI can generate damped and delayed responses of both durable and nondurable consumption to income shocks.

Our paper is related to two strands of the literature. First, the LQ approximation method in step 1 of our solution approach is related to the Hamiltonian approach to the deterministic continuous-time control problem in Magill (1977). Judd (1998) points out that it may deliver an inaccurate solution if one adopts an ad hoc LQ approximation method by simply computing a quadratic approximation to the objective function and a first-order approximation to the model structural relations. Magill's approach has been applied to the optimal policy problems in discrete time by Levine et al. (2008) and Benigno and Woodford (2012). We extend Magill's approach to incorporate partial information in discrete time.

Second, our paper is related to the RI literature in the LQG framework.[2] Most papers in this literature focus on dynamic tracking problems in which states follow exogenous dynamics. For these problems, Sims (2003) proposes a brute force optimization method for the univariate case in the frequency domain.[3] Peng (2005), Peng and Xiong (2006), and Maćkowiak and Wiederholt (2009) propose methods under a signal independence assumption. For a general case with one action that is driven by possibly multiple autoregressive moving average (ARMA) processes, Maćkowiak et al. (2018) develop a method based on the state space representation without any ad hoc restriction on the signal form. Afrouzi and Yang (2021) develop numerical methods to solve general multivariate tracking problems.

While Sims (2003, 2011) formulates rationally inattentive control problems with endogenous state dynamics in the LQG framework, his solution approach applies to the univariate case in which the optimal signal is equal to the state plus a noise. Luo (2008) and Luo et al. (2015) apply Sims's approach when a multivariate consumption/saving problem can be reduced to the univariate case. As aforementioned, MWY (2022) propose an approach to study general multivariate rationally inattentive control problems. Unlike other approaches in the literature that only compute the steady state solution for the optimal information structure with the subjective discount factor equal to 1, both MWY (2022) and Afrouzi and Yang (2021) compute transition dynamics as well as steady state with any discount factor between 0 and 1. For simplicity here, we focus on the steady state information structure only.

For nonlinear control or tracking problems, the literature typically uses LQ approximations to transform these problems into tracking problems in the LQG framework. While this procedure is feasible in some cases [e.g. Maćkowiak and Wiederholt (2009, 2015, 2020), Maćkowiak et al. (2018), and Zorn (2018)], it is cumbersome and may not apply to more complicated control problems, in which controls affect state dynamics. For these control problems, ad hoc LQ approximations may lead to inaccurate solutions [Judd (1998)]. MWY (2022) apply the Kydland and

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

Prescott (1982) approach to conduct LQ approximations to a nonlinear investment problem. This approach is a special case of our new approach in this paper with linear constraints only. Our new approach can apply to a much wider class of economic problems beyond those in MWY (2022) and Maćkowiak–Wiederholt's studies. Moreover, the Matlab toolbox implements our approach and should be useful for researchers.

The remainder of the paper proceeds as follows. Section 2 formulates a general rationally inattentive control problem. Section 3 presents an LQ approximation approach. Section 4 describes a Matlab toolbox to implement our LQ approximation approach. Section 5 studies five examples to demonstrate the toolbox. Section 6 concludes. All proofs are relegated to an appendix.

2. Rationally inattentive control problem

In this section, we first present the standard control model under full information and then formulate the rationally inattentive control problem.

Consider an infinite-horizon discrete-time setup and time is denoted by $t \geq 0$. Let x_t denote an $n_x \times 1$ vector of states at time t. States evolve according to the dynamics:

$$x_{t+1} = g(x_t, u_t, \epsilon_{t+1}), \ t \geq 0, \tag{1}$$

where x_0 is exogenously given, u_t is an $n_u \times 1$ control vector, ϵ_{t+1} is an $n_\epsilon \times 1$ white noise vector with an identity covariance matrix, and $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\epsilon} \to \mathbb{R}^{n_x}$. The state vector x_t may consist of exogenous components such as AR(1) shocks and endogenous components such as capital.

Under full information, the DM observes the history of shocks $\epsilon^t = \{\epsilon_1, \ldots, \epsilon_t\}$ at any time $t \geq 1$ and the history of states $x^t = \{x_0, x_1, \ldots, x_t\}$. A control plan $\{u_t\}$ is a sequence of controls that map x^t to $u_t(x^t)$ for $t \geq 0$, where all functions $u_t(x^t)$ are measurable. Let Γ_0 denote the set of all such control plans. Then the DM's objective under full information is to solve the following problem:

$$\sup_{\{u_t\} \in \Gamma_0} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t f(x_t, u_t)\right]$$

subject to (1), where $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ and $\beta \in (0, 1)$.

Next, we turn to the control problem under RI. Suppose that the initial state x_0 is random and has a probability measure denoted by $\mu_0(dx_0)$. The DM does not fully observe the state x_t at any time $t \geq 0$. The DM can acquire endogenous information about the state by paying information costs. Assume that information costs are measured in utility units. The DM chooses a signal about the state x_t with realization s_t in some signal space $\mathbb{S}$. The DM's choice is a strategy pair $(\{u_t\}, \{q_t\})$ composed of

- an information strategy $\{q_t\}$ consisting of a sequence of distributions $q_t(ds_t | x^t, s^{t-1})$ for all $s^t, x^t, t \geq 0, s^{-1} = \varnothing$;
- a control plan $\{u_t\}$ consisting of a sequence of functions $u_t : \mathbb{S}^t \to \mathbb{R}^{n_u}$, which deliver a control $u_t = u_t(s^t)$ after observing a history of signals s^t for $t \geq 0$.

Let Γ denote the set of all such strategies $(\{u_t\}, \{q_t\})$. Following Sims (2011), we model the information cost by the discounted mutual information. To define it formally, we need to construct the joint distribution of states and signals. The function g and the distribution of ϵ_{t+1} induce a transition kernel for the state, denoted by $\pi(dx_{t+1} | x_t, u_t) = \Pr(g(x_t, u_t, \epsilon_{t+1}) \in dx_{t+1} | x_t, u_t)$.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

The state transition kernel π and the strategy $(\{u_t\}, \{q_t\})$ induce a sequence of joint distributions for x^{t+1} and s^t recursively

$$\mu_{t+1}(dx^{t+1}, ds^t) = \pi(dx_{t+1}|x_t, u_t(s^t))\, q_t(ds_t|x^t, s^{t-1})\, \mu_t(dx^t, ds^{t-1}),$$

where $\mu_0(dx^0, ds^{-1}) = \mu_0(dx_0)$ is given and $s^{-1} = \varnothing$. Using this sequence of distributions, we can compute the prior/predictive distributions $\mu_t(dx_t|s^{t-1})$ and the posteriors $\mu_t(dx_t|s^t)$, and hence we can define the discounted information cost as:

$$\sum_{t=0}^{T} \beta^t \mathcal{I}(x_t; s_t|s^{t-1}), \tag{2}$$

where

$$\mathcal{I}(x_t; s_t|s^{t-1}) = H(x_t|s^{t-1}) - H(x_t|s^t). \tag{3}$$

Here, $H(X|Y)$ denotes the conditional entropy of a random variable X given Y.[4] Entropy measures the amount of uncertainty. Equation (3) shows that the mutual information measures the reduction of uncertainty about the state after observing additional information. As Sims (2011) and MWY (2022) argue, introducing discounting in equation (2) ensures dynamic consistency in the choice of optimal information structure and thus one can apply the dynamic programming method.

Now we are ready to formulate the rationally inattentive control problem as follows:

Problem 1. *(Rationally inattentive control)*

$$\sup_{(\{u_t\},\{q_t\})\in\Gamma} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t f(x_t, u_t)\right] - \lambda \sum_{t=0}^{\infty} \beta^t \mathcal{I}(x_t; s_t|s^{t-1}) \tag{4}$$

subject to equation (1)*, where* $\lambda > 0$.

The parameter $\lambda > 0$ transforms the discounted mutual information into utility units. If $\lambda = 0$, then information is free and thus Problem 1 is reduced to the standard control problem under full information. If $\lambda > 0$, acquiring information incurs information costs measured in utility units and thus reduces the DM's utility. The parameter λ may be interpreted as the Lagrange multiplier associated with an information processing constraint. Then, Problem 1 can be interpreted as a relaxed problem derived from a constrained optimization problem given information processing constraints [Sims (2011) and MWY (2022)]. Without additional structure, Problem 1 is hard to analyze as one has to solve for both the optimal control $\{u_t\}$ and the optimal information structure $\{q_t\}$. In the next section, we propose an LQG approximation approach.

3. Linear quadratic approximation

We first fix the information structure and solve the optimal control problem under partial information. Specifically, we take the filtration $\{s^t\}$ generated by the histories of signals s^t as given. The DM's information set is $\{s^t\}$. Let Γ_1 denote the set of control plans that are adapted to the filtration $\{s^t\}$.

Consider the following control problem under partial information:

$$\sup_{\{u_t\}\in\Gamma_1} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t f(x_t, u_t)\right]$$

subject to equation (1). Under the standard concavity and differentiability conditions on f and g, the following first-order conditions are necessary and sufficient for optimality:[5]

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

$$u_t:\ \mathbb{E}_t\left[f_u(x_t,u_t)\right]+\Lambda_t'\mathbb{E}_t\left[g_u(x_t,u_t,\epsilon_{t+1})\right]=0, \tag{5}$$

$$x_{t+1}:\ \Lambda_t=\beta\mathbb{E}_t\left[f_x(x_{t+1},u_{t+1})+\Lambda_{t+1}'g_x(x_{t+1},u_{t+1},\epsilon_{t+2})\right], \tag{6}$$

where $\beta^t\Lambda_t$ is the Lagrange multiplier associated with equation (1) and $\mathbb{E}_t[\cdot]\equiv\mathbb{E}\left[\cdot|s^t\right]$. Here, $\{\Lambda_t\}$ is adapted to the filtration $\left\{s^t\right\}$. Unlike in the full information case, the state x_t is not observable and thus equation (5) involves estimate $\mathbb{E}_t[\cdot]\equiv\mathbb{E}\left[\cdot|s^t\right]$ given partial information s^t.

In a nonstochastic steady state, $\epsilon_t=0$, $x_t=\bar{x}$, $u_t=\bar{u}$, and $\Lambda_t=\overline{\Lambda}$ for all t. It follows from equations (5) and (6) that

$$f_u(\bar{x},\bar{u})+\overline{\Lambda}'g_u(\bar{x},\bar{u},0)=0, \tag{7}$$

$$\beta f_x(\bar{x},\bar{u})+\beta\overline{\Lambda}'g_x(\bar{x},\bar{u},0)=\overline{\Lambda}. \tag{8}$$

Together with the steady state version of equation (1),

$$\bar{x}=g(\bar{x},\bar{u},0),$$

equations (7) and (8) determine a steady state solution for $\left(\bar{x},\bar{u},\overline{\Lambda}\right)$. Suppose that a solution exists. Then linearizing equations (1), (5), and (6) around this steady state yields the system:

$$g_u'\widetilde{\Lambda}_t+\left(f_{uu}+\overline{\Lambda}'g_{uu}\right)\widetilde{u}_t+\left(f_{ux}+\overline{\Lambda}'g_{ux}\right)\mathbb{E}_t[\widetilde{x}_t]=0, \tag{9}$$

$$\widetilde{\Lambda}_t=\beta\mathbb{E}_t\left[\left(f_{xx}+\overline{\Lambda}'g_{xx}\right)\widetilde{x}_{t+1}+\left(f_{xu}+\overline{\Lambda}'g_{uu}\right)\widetilde{u}_{t+1}+g_x'\widetilde{\Lambda}_{t+1}\right], \tag{10}$$

$$\widetilde{x}_{t+1}=g_x\widetilde{x}_t+g_u\widetilde{u}_t+g_\epsilon\epsilon_{t+1}, \tag{11}$$

where all partial derivatives are evaluated at the nonstochastic steady state $(\bar{x},\bar{u},0)$ and a variable with a tilde denotes the level deviation from its steady state, for example, $\widetilde{x}_t\equiv x_t-\bar{x}$. As the state x_t or $\widetilde{x}_t$ is unobservable, the DM must estimate it given information s^t. We thus have $\mathbb{E}_t[\widetilde{x}_t]$ in equation (9).

The above linear system can be solved by the standard method, and the solution takes a linear form. In particular, the certainty equivalence principle holds in that the optimal policy satisfies

$$\widetilde{u}_t=-F\mathbb{E}_t[\widetilde{x}_t], \tag{12}$$

where F is the same as that obtained in the deterministic linear system with $\epsilon_t=0$ for all t.

Next, we study how the optimal information structure is determined. To apply the LQG framework of MWY (2022), we adopt the LQ approximation approach adapted from Magill (1977) and take into account the impact of partial information. This approach delivers a quadratic approximation of the objective function and a linear approximation of the constraint. To ensure the resulting LQ control problem gives a linear solution that is the same as equation (12), it is critical to approximate the Hamiltonian function defined as:

$$H(x_t,u_t,\epsilon_{t+1})=f(x_t,u_t)+\overline{\Lambda}'g(x_t,u_t,\epsilon_{t+1}). \tag{13}$$

Notice that $\overline{\Lambda}$ in H is the steady state Lagrange multiplier. Then, we have the following result:

Lemma 1. *Suppose that f and g are twice continuously differentiable and $\lim_{T\to\infty}\mathbb{E}[\widetilde{x}_T]=0$. Then given equation (1),*

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t f(x_t,u_t)\right]\approx\frac{1}{2}\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\widetilde{x}_t',\widetilde{u}_t'\right]\begin{bmatrix}H_{xx}&H_{xu}\\H_{ux}&H_{uu}\end{bmatrix}\begin{bmatrix}\widetilde{x}_t\\\widetilde{u}_t\end{bmatrix}$$
$$+\frac{f(\bar{x},\bar{u})}{1-\beta}+\frac{1}{2}Tr(H_{\epsilon\epsilon}),$$

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

up to second-order moments of $(\widetilde{x}_t, \widetilde{u}_t, \epsilon_{t+1})$ *, where* $Tr(\cdot)$ *denotes a trace operator, and* H_{xx}, H_{xu}, H_{ux}, *and* H_{uu} *denote the second-order partial derivatives of* H *evaluated at the nonstochastic steady state* $(\overline{x}, \overline{u}, 0)$.

Ignoring the constant and higher-order terms, the following lemma shows that the LQ problem gives the same solution as that obtained by linearizing the first-order conditions for the original nonlinear control problem.

Lemma 2. *The following problem*

$$\max_{\{\widetilde{u}_t\}\in\Gamma_1} \frac{1}{2}\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\widetilde{x}_t', \widetilde{u}_t'\right]\begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu}\end{bmatrix}\begin{bmatrix}\widetilde{x}_t \\ \widetilde{u}_t\end{bmatrix}$$

subject to equation (11) gives the same solution as that delivered by the system (9), (10), and (11).

The LQ approach of Kydland and Prescott (1982) is a special case. They assume that the constraint function g is linear. Then the Hessian matrix of H evaluated at the steady state is the same as the Hessian matrix of the objective function f evaluated at the steady state. Thus, taking quadratic approximation of H is the same as that of f. While one can often use a suitable change of variables to make the constraint function linear, the Hamiltonian approach is more general and more convenient.

Notice that the ad hoc LQ approach differs from our correct Hamiltonian approach by the second-order derivative terms of function g. The intuition comes from the linearized first-order conditions (9) and (10). The ad hoc LQ approach disregards second-order derivative terms of g in equations (9) and (10), leading to an inaccurate solution if g is nonlinear.

By Lemma 2, we obtain the following LQ approximation of the rationally inattentive control problem.

Problem 2. *(Rationally inattentive LQ control)*

$$\sup_{(\{\widetilde{u}_t\},\{q_t\})\in\Gamma} \frac{1}{2}\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\widetilde{x}_t', \widetilde{u}_t'\right]\begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu}\end{bmatrix}\begin{bmatrix}\widetilde{x}_t \\ \widetilde{u}_t\end{bmatrix} - \lambda\sum_{t=0}^{\infty}\beta^t\mathcal{I}\left(x_t; s_t|s^{t-1}\right)$$

subject to equation (11).

To simplify the computation of the mutual information and stay in the LQG framework of Sims (2011) and MWY (2022), we consider Gaussian shocks only.

Assumption 1. *(i) The initial state* x_0 *is Gaussian with mean* $\overline{x}_0$ *and covariance matrix* Σ_{-1}. *(ii) The innovation* ϵ_t *is identically and independently drawn from a standard Gaussian distribution.*

Problem 2 corresponds to the infinite-horizon version of Problem 2 in MWY (2022) using the following notations:

$$Q \equiv -\frac{1}{2}H_{xx},\ R \equiv -\frac{1}{2}H_{uu},\ S \equiv -\frac{1}{2}H_{xu}, \tag{14}$$

$$A \equiv g_x,\ B \equiv g_u,\ W \equiv g_\epsilon g_\epsilon'. \tag{15}$$

As shown in Sims (2011) and Tanaka et al. (2017), the optimal information structure will be Gaussian in that the optimal estimate of $\widetilde{x}_t$ conditional on s^t is Gaussian with mean zero and covariance matrix $\Sigma_t = \mathbb{E}\left[\left(\widetilde{x}_t - \mathbb{E}\left[\widetilde{x}_t|s^t\right]\right)\left(\widetilde{x}_t - \mathbb{E}\left[\widetilde{x}_t|s^t\right]\right)'|s^t\right]$. Moreover, a linear signal s_t of the form

$$s_t = C_t\widetilde{x}_t + v_t \tag{16}$$

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

can generate such a posterior covariance matrix Σ_t, where v_t is a Gaussian white noise with mean zero and covariance matrix V_t. This noise is independent of $\{\epsilon_t\}$. As Σ_t will be endogenously chosen, and both C_t and V_t are also endogenous and satisfy

$$C_t' V_t^{-1} C_t = \Sigma_t^{-1} - (A\Sigma_{t-1}A + W)^{-1}.$$

Given Gaussian uncertainty, the mutual information takes an explicit form:

$$\mathcal{I}\left(x_0; s_0 | s^{-1}\right) = \mathcal{I}\left(\widetilde{x}_t; s_t | s^{t-1}\right) = \frac{1}{2}\log\det\left(\Sigma_{-1}\right) - \frac{1}{2}\log\det(\Sigma_0), \tag{17}$$

$$\mathcal{I}\left(x_t; s_t | s^{t-1}\right) = \mathcal{I}\left(\widetilde{x}_t; s_t | s^{t-1}\right) = \frac{1}{2}\log\det\left(A\Sigma_{t-1}A' + W\right) - \frac{1}{2}\log\det\Sigma_t, \tag{18}$$

for $t \geq 1$.

Now, we have mapped Problem 2 into the framework of MWY (2022). We can then apply their method and their toolbox described in Miao and Wu (2021) to solve this problem. In the next section, we describe the details.

4. A Matlab toolbox

In this section, we describe a Matlab toolbox to implement the LQG approximation approach described in the previous section. The only inputs that the user needs to provide are the state vector x_t, the control vector u_t, the objective function f, the constraint function g, and the non-stochastic steady state values of (x_t, u_t, Λ_t) given by $\left(\overline{x}, \overline{u}, \overline{\Lambda}\right)$. For the example in Section 5.2, the Matlab code RBCPref_model.m describes x_t, u_t, f, and g, and the code RBCPref_model_ss.m solves the steady state. We use the Matlab Symbolic Math toolbox to compute the analytical derivatives of f and g in the code anal_derivative.m. The code num_derivative.m evaluates the analytical derivatives at the steady state values.

The code RBCPref_Run.m is the main code that implements the following steps.

Step 1. Compute matrices Q, R, S, A, B, W in equations (14) and (15). Then use the code LQG.m to compute the Riccati equation for P:

$$P = Q + \beta A'PA - \left(\beta A'PB + S\right)\left(R + \beta B'PB\right)^{-1}\left(\beta B'PA + S'\right), \tag{19}$$

and

$$F = \left(R + \beta B'PB\right)^{-1}\left(S' + \beta B'PA\right).$$

The optimal policy is given by equation (12).

Step 2. The optimal information structure $\{\Sigma_t\}_{t=0}^{\infty}$ solves the following problem.

Problem 3. *(Optimal information structure for Problem 2)*

$$\min_{\{\Sigma_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\mathrm{tr}(\Omega\Sigma_t) + \lambda\mathcal{I}\left(x_t; s_t | s^{t-1}\right)\right]$$

subject to equations (17) and (18),

$$\Sigma_t \preceq A\Sigma_{t-1}A' + W, \tag{20}$$

$$\Sigma_0 \preceq \Sigma_{-1}, \tag{21}$$

for $t \geq 1$,[6] *where* Ω *is given by:*

$$\Omega = F'(R + \beta B'PB)F. \tag{22}$$

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

For this problem to be a well-defined convex optimization problem, we make the following assumption [Afrouzi and Yang (2021) and MWY (2022)]:

Assumption 2. *$W \succeq 0$ and $AA' + W \succ 0$.*

If this assumption is violated, then the matrix $A\Sigma_{t-1}A' + W$ may be singular even though Σ_{t-1} is nonsingular, for example, $\Sigma_{t-1} = I$. In this case, the mutual information in equation (18) is not well defined. Moreover, this assumption is also sufficient for the convexity of Problem 3. We will focus on the steady state solution Σ to the above problem. The Matlab code RI_SS_FOC.m computes this solution.

Step 3. Use the code RI_SIG.m to compute the steady state optimal signal structure (C, V) that generates the optimal steady state posterior covariance matrix Σ. Notice that the optimal signal structure (C, V) is derived from the following equation:

$$C'V^{-1}C = \Sigma^{-1} - (A\Sigma A + W)^{-1},$$

and the solution is not unique.

Step 4. Use the code RI_IRF3.m to generate impulse response functions (IRFs) using the steady state Kalman filter for the state space system:

$$\begin{aligned} \widetilde{x}_{t+1} &= A\widetilde{x}_t + B\widetilde{u}_t + L\epsilon_{t+1}, \\ \widetilde{u}_t &= -F\mathbb{E}_t\left[\widetilde{x}_t\right], \\ s_t &= C\widetilde{x}_t + v_t, \end{aligned} \tag{23}$$

where $L \equiv g_\epsilon$ and v_t is a Gaussian white noise with covariance matrix V and is independent of $\{\epsilon_t\}$. The Kalman filter is given by:

$$\widehat{x}_t = (I - KC)(A - BF)\widehat{x}_{t-1} + K(C\widetilde{x}_t + v_t), \tag{24}$$

$$\widetilde{x}_{t+1} = A\widetilde{x}_t - BF\widehat{x}_t + L\epsilon_{t+1},\ t \geq 0, \tag{25}$$

where $\widehat{x}_t = \mathbb{E}_t\left[\widetilde{x}_t\right]$ and the matrix K is the Kalman gain:

$$K \equiv \left(A\Sigma A' + W\right)C'\left[C\left(A\Sigma A' + W\right)C' + V\right]^{-1}. \tag{26}$$

So far we have described how to solve a linear approximate solution. In economics, one is often interested in a log-linear approximate solution for positive variables. Such a solution can be easily found by noting that

$$\log x_t - \log \overline{x} \approx \frac{\widetilde{x}_t}{\overline{x}}, \tag{27}$$

for any positive variable x_t up to a first-order approximation. We can then divide the linear approximate solution for any variable of interest by its steady state value to obtain its log-linear approximate solution. The other way is to make a change of variables $x_t^l = \log x_t$ and $u_t^l = \log u_t$ and linearize with respect to x_t^l and u_t^l. We also replace the signal equation (16) by:

$$s_t = C_t\widetilde{x}_t^l + v_t$$

and replace the mutual information $\mathcal{I}\left(\widetilde{x}_t; s_t|s^{t-1}\right)$ by $\mathcal{I}\left(\widetilde{x}_t^l; s_t|s^{t-1}\right)$. We can then apply the previous toolbox.

5. Examples

In this section, we present five examples to illustrate our toolbox. The first example illustrates some pitfalls of the ad hoc LQ approximation method. The next three examples are about social planner

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

problems as in the RBC literature. We do not consider decentralized market equilibrium because our toolbox applies only to decision problems. The final example studies a consumption/saving problem and shows that one can suitably define the state and control variables to transform a complicated decision problem into our framework. For all examples, we focus on the solution methods instead of quantitative economic implications, and thus parameter values are chosen for illustration, but not for matching data closely.

5.1. Pitfalls

Judd (1998) uses a simple deterministic growth model in continuous time to show that the ad hoc LQ approximation method can generate inaccurate solutions. In this subsection, we present a stochastic growth model in discrete time to highlight additional issues that may arise under full information and under limited information with RI.

Formally, consider a planner's choice of consumption and capital processes $\{C_t\}$ and $\{K_{t+1}\}$ under full information:

$$\max \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \log C_t\right]$$

subject to

$$C_t + K_{t+1} = \exp(a_{1t} + a_{2t})\, K_t^{\alpha}, \quad (\Lambda_{3t})$$

where a_{1t} and a_{2t} follow independent AR(1) processes:

$$a_{i,t+1} = \rho_i a_{it} + \epsilon_{i,t+1}, \quad (\Lambda_{it}), \quad i = 1, 2. \tag{28}$$

Here, $\epsilon_{i,t}$, $i = 1, 2$, are Gaussian white noises with variances σ_i^2. Let Λ_{it}, $i = 1, 2, 3$, be the Lagrange multipliers associated with the above three equations.

The above control problem admits a closed-form solution:

$$K_{t+1} = \alpha\beta \exp(a_{1t} + a_{2t})\, K_t^{\alpha}, \; C_t = (1 - \alpha\beta) \exp(a_{1t} + a_{2t})\, K_t^{\alpha}.$$

The deterministic steady state is given by:

$$\overline{K} = (\alpha\beta)^{1/(1-\alpha)}, \; \overline{C} = (1 - \alpha\beta)\, \overline{K}^{\alpha},$$

$$\overline{\Lambda}_3 = \frac{1}{\overline{C}}, \; \overline{\Lambda}_1 = \frac{\beta \overline{\Lambda}_3 \overline{K}^{\alpha}}{1 - \beta\rho_1}, \; \overline{\Lambda}_2 = \frac{\beta \overline{\Lambda}_3 \overline{K}^{\alpha}}{1 - \beta\rho_2}.$$

The true linearized solution around the steady state is given by:

$$\begin{aligned} \widetilde{K}_{t+1} &= \alpha \widetilde{K}_t + \alpha\beta \overline{K}^{\alpha} a_{1t} + \alpha\beta \overline{K}^{\alpha} a_{2t}, \\ \widetilde{C}_t &= (1 - \alpha\beta)\, \alpha \overline{K}^{\alpha-1} \widetilde{K}_t + (1 - \alpha\beta)\, \overline{K}^{\alpha} a_{1t} + (1 - \alpha\beta)\, \overline{K}^{\alpha} a_{2t}. \end{aligned} \tag{29}$$

Then both optimal consumption and capital are stationary processes.

If one adopts the ad hoc LQ approximation method, one solves the following approximated problem:

$$\max_{\{\widetilde{C}_t, \widetilde{K}_{t+1}\}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \left(\log \overline{C} + \frac{1}{\overline{C}} \widetilde{C}_t - \frac{1}{2\overline{C}^2} \widetilde{C}_t^2\right)\right]$$

subject to the linearized constraints:

$$\widetilde{C}_t + \widetilde{K}_{t+1} = \overline{K}^{\alpha} (a_{1t} + a_{2t}) + \alpha \overline{K}^{\alpha-1} \widetilde{K}_t.$$

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

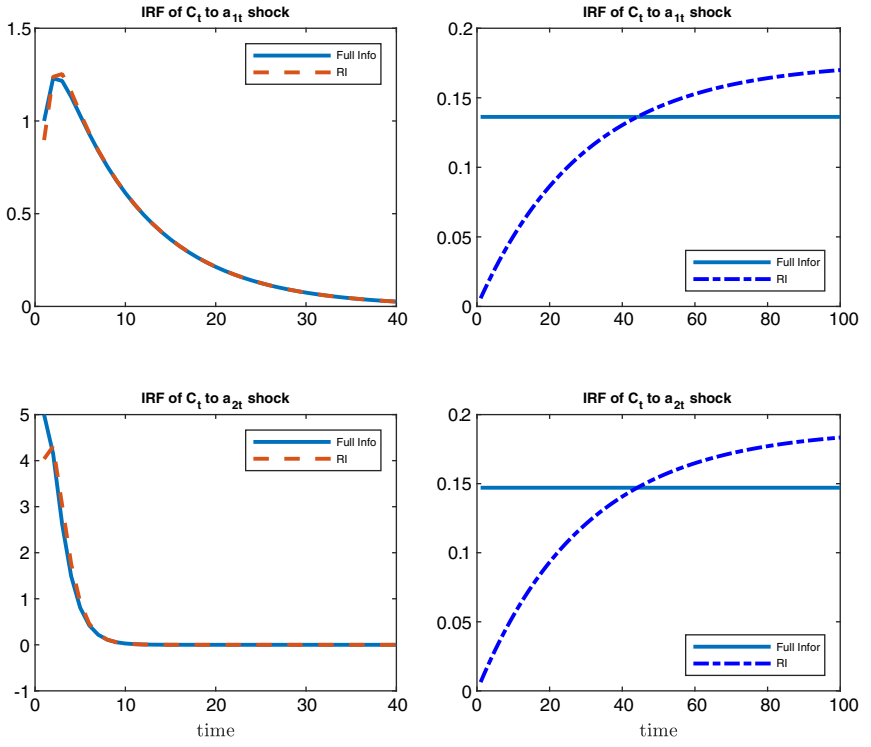


Figure 1. Impulse responses of C_t to a one-standard-deviation innovation in the two TFP components under full information and under RI. The left and right two panels display results using our method and the ad hoc LQ approximation method, respectively. All vertical axes are measured in percentage changes from the steady state.

This problem becomes a permanent income model of consumption in the LQ framework. The Euler equation is given by:

$$\widetilde{C}_t = \mathbb{E}_t \widetilde{C}_{t+1},$$

which implies that optimal consumption is nonstationary and follows a random walk. It follows that the optimal consumption level is proportional to the permanent income:

$$\widetilde{C}_t = \left(\alpha \overline{K}^{\alpha-1} - 1\right) \widetilde{K}_t + \frac{\left(\alpha \overline{K}^{\alpha-1} - 1\right) \overline{K}^{\alpha}}{\alpha \overline{K}^{\alpha-1} - \rho_1} a_{1t} + \frac{\left(\alpha \overline{K}^{\alpha-1} - 1\right) \overline{K}^{\alpha}}{\alpha \overline{K}^{\alpha-1} - \rho_2} a_{2t}.$$

Clearly, this solution is inaccurate compared to equation (29).

Now we turn to the case under limited information. Suppose that the planner does not fully observe any state of the model and can acquire information about the states subject to discounted information costs. Since our method nests the full information case, it delivers the same linearized solution as in equation (29). To show quantitative implications, we set parameter values $\alpha = 0.33$, $\beta = 0.99$, $\rho_1 = 0.9$, $\sigma_1 = 0.01$, $\rho_2 = 0.5$, $\sigma_2 = 0.05$, and $\lambda = 0.005$. Figure 1 plots the IRFs of consumption to a one-standard deviation innovation shock to the two total factor productivity (TFP) components. The left two panels display the result derived from our method and toolbox. One can verify that the numerical solution under full information is consistent with the analytical solution in equation (29). Under RI, the initial consumption response is weaker, and the later response is delayed. It takes a longer time for consumption to return to its steady state level.

The right two panels display the result derived from the ad hoc LQ approximation method. Optimal consumption under full information (solid lines) jumps to a higher steady state level

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

immediately and stays at that level forever. But the initial consumption response under RI is weaker. Consumption gradually rises to a higher steady state level. This result is consistent with that in Luo (2008) and MWY (2022) for permanent income LQ models of consumption but is drastically different from the true linearized solution shown earlier.

The above example demonstrates that the ad hoc LQ approximation method may generate not only an inaccurate solution but also qualitatively different economic behavior both under full information and under limited information with RI. The ad hoc LQ approximation method works only when the constraints (1) are linear [see Judd (1998)]. In this case, the Hessian for the term $\overline{\Lambda}' g(x_t, u_t, \epsilon_{t+1})$ in equation (13) is equal to zero. This case can happen for pure tracking problems when exogenous states follow AR(1) processes like equation (28).

It merits emphasis that our LQ approach implies that the certainty equivalence principle holds for the approximated LQ control problem. This implies that uncertainty does not matter for the linear decision rule. Because this principle does not apply to the general nonlinear control problem, our LQ approach is not well suited to handle questions such as welfare comparisons across alternative stochastic environments. For example, Kim and Kim (2003) show that in a simple two-agent economy, a welfare comparison based on an evaluation of the utility function using a linear approximation to the policy function may yield the spurious result that welfare is higher under autarky than under full risk sharing. Our LQ approach is also not suitable to study questions related to risk premium, in which uncertainty plays an important role. To study these questions, one has to take at least a second-order approximation to the decision rule. In this case, our solution method under RI does not apply.

5.2. Planner's problem I: preference shock

We next consider a more complicated social planner's problem. Under full information, the planner's objective is to maximize the representative household's utility over consumption $\{C_t\}$ and labor $\{N_t\}$:

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t U(C_t, N_t, z_t)\right],\ U(C, N, z) = \exp(z)\log(C) - \chi\frac{N^{1+\nu}}{1+\nu},$$

subject to

$$C_t + I_t = \exp(a_t)\,(e_t K_t)^{\alpha} N_t^{1-\alpha}, \tag{30}$$

$$K_{t+1} = (1 - \delta(e_t))\,K_t + I_t, \tag{31}$$

$$z_{t+1} = \rho_z z_t + \epsilon_{z,t+1}, (\Lambda_{1t}) \tag{32}$$

$$a_{t+1} = \rho_a a_t + \epsilon_{a,t+1},\ \ (\Lambda_{2t}) \tag{33}$$

for $t \geq 0$, where K_t, I_t, e_t, a_t, and z_t represent capital, investment, capital utilization rate, TFP shock, and preference shock, respectively. Equations (30) and (31) are the resource constraint and the law of motion for capital, respectively. Equations (32) and (33) give AR(1) process specifications, where $\epsilon_{a,t+1}$ and $\epsilon_{z,t+1}$ are independent Gaussian white noises with variances σ_a^2 and σ_z^2.

More intensively utilized capital raises capital efficiency but also makes capital depreciate faster. Let the depreciation rate satisfy

$$\delta(e_t) = \delta + \phi_e\frac{(e_t)^{\gamma} - 1}{\gamma},\ \gamma > 1,\ \phi_e > 0.$$

Suppose that the planner does not fully observe any state of the model and can acquire information about the states subject to discounted information costs. We choose the state vector as

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

$x_t = (z_t, a_t, K_t)'$ and the control vector as $u_t = (C_t, N_t, e_t)'$. Assume that the initial values of the states K_t, a_t, and z_t are drawn from independent Gaussian distributions. After using equation (30) to substitute I_t in equation (31), we obtain the state transition equation:

$$K_{t+1} = (1-\delta)\, K_t + \exp(a_t)\, (e_t K_t)^{\alpha}\, N_t^{1-\alpha} - C_t. \quad (\Lambda_{3t}) \tag{34}$$

Then the planner's problem is transformed into our framework.

Let Λ_{it}, $i = 1, 2, 3$, denote the (undiscounted) Lagrange multipliers associated with equations (32), (33), and (34). Their steady state values are given by:

$$\overline{\Lambda}_1 = \frac{\beta \log\left(\overline{C}\right)}{1-\beta\rho_z},\ \overline{\Lambda}_2 = \frac{\beta U_c\left(\overline{C}, \overline{N}\right) \overline{K}^{\alpha} \overline{N}^{1-\alpha}}{1-\beta\rho_a},\ \overline{\Lambda}_3 = U_c\left(\overline{C}, \overline{N}\right).$$

To solve this problem numerically, we set parameter values as $\alpha = 0.33$, $\delta = 0.025$, $\nu = 1$, $\beta = 0.99$, $\gamma = 1.2$, $\rho_a = 0.95$, $\rho_z = 0.8$, $\sigma_a = \sigma_z = 0.01$, and $\lambda = 0.002$. We also choose $\chi = 7.8827$ such that the steady state labor is $1/3$ and $\phi_e = 0.0351$ such that the steady state capital utilization rate is 1. The linearized solution gives the policy function $\widetilde{u}_t = -F\mathbb{E}\left[\widetilde{x}_t | s^t\right]$, where

$$F = \begin{bmatrix} -0.6978 & -0.3520 & -0.0213 \\ -0.0286 & -0.2855 & 0.0061 \\ -0.0662 & -1.8091 & 0.0955 \end{bmatrix}.$$

Notice that the Kydland–Prescott LQ approximation method does not apply to the above formulation because equation (34) is nonlinear.

One can easily check that the linearized state transition matrix A is invertible, but the covariance matrix W for the state transition noise is not invertible. Assumption 2 is satisfied and one can apply our toolbox to solve the RI problem. The optimal steady state posterior covariance matrix is given by:

$$\Sigma = \begin{bmatrix} 0.0276 & -0.0028 & -0.0286 \\ -0.0028 & 0.0348 & 0.0382 \\ -0.0286 & 0.0382 & 1.9343 \end{bmatrix} \times 10^{-2}.$$

The posterior variances of z_t and a_t are smaller than their unconditional values 5.2632×10^{-4} and 0.001. Even though z_t and a_t are ex ante independent, they are ex post negatively correlated. This is due to the following optimal signal form computed from our toolbox:

$$s_t = 0.3015 z_t + 0.9507 a_t + 0.0725 \widetilde{K}_t + \nu_t, \tag{35}$$

where ν_t is a Gaussian white noise with variance 0.0023. When observing the same realization of the signal s_t, the planner may attribute a positive preference shock z_t to a negative TFP shock a_t, holding $\widetilde{K}_t$ and ν_t constant. On the other hand, observing a positive signal, the planner may attribute it to a positive TFP shock or a positive preference shock, holding $\widetilde{K}_t$ and ν_t constant.

We find both methods discussed in Section 4 generate the same log-linearized optimal policy functions and almost the same IRFs for the log-linearized solution. Here, we present the IRFs using the first method. As is well known, there is no comovement in response to a preference shock under full information. In this case, a positive preference shock causes consumption and labor to rise, but investment to fall, as shown in Figure 2. By contrast, under RI, the optimal signal form in equation (35) implies that the inattentive planner confuses a preference shock with a TFP shock given the same realization of the signal. Observing a positive signal, the planner may interpret a positive preference shock as a positive TFP shock. Given the parameterization, consumption, labor, investment, and hence output all rise on impact in response to a positive preference shock.

Notice that capital utilization plays an important role in this comovement result. By results not reported here but available upon request, we find that there is no comovement even under RI if

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

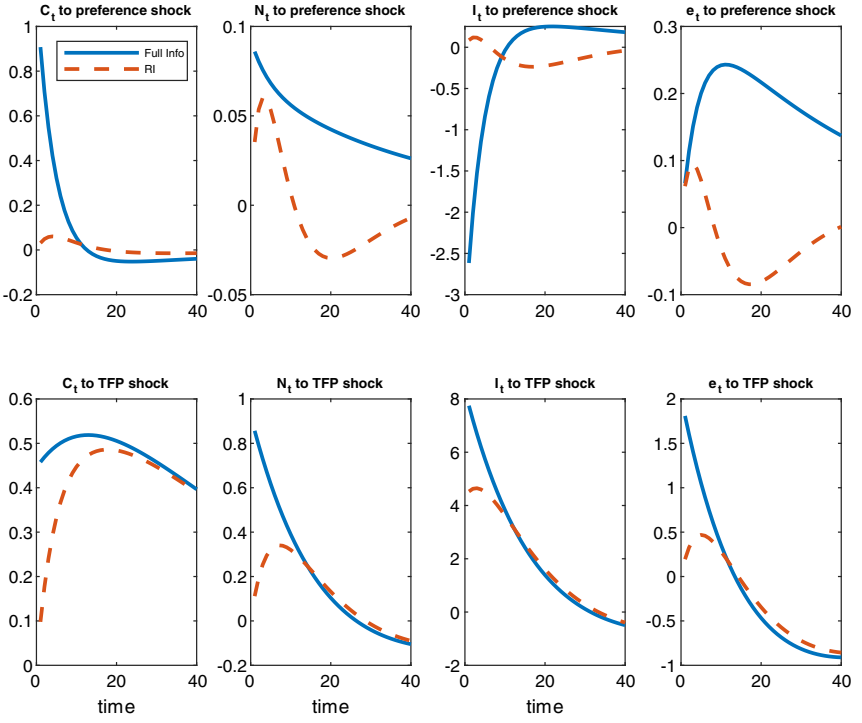


Figure 2. Impulse responses of I_t, N_t, and C_t to a one-percent innovation in the TFP and preference shocks under full information and under RI. All vertical axes are measured in percentage changes from the steady state.

there is no endogenous capital utilization. The intuition is as follows: a positive preference shock raises consumption and labor. But the magnitude of the rise of labor and hence output is too small, leading investment to fall by the resource constraint. With endogenous capital utilization, a rise in labor also causes capital to be utilized more intensively and thus raises output further. The large increase in output allows investment to rise.[7]

The bottom panels of Figure 2 shows that RI generates damped and delayed responses to TFP shocks. Moreover, the responses are hump-shaped, even though there is no adjustment cost in this model.

As shown in MWY (2022), the signal dimension weakly increases as the information cost parameter λ decreases. Intuitively, as the information cost becomes smaller, the planner acquires more information. For example, when $\lambda = 0.0005$, the optimal signal takes the form:

$$s_t = \begin{bmatrix} -0.3704 & -0.9285 & -0.0249 \\ 0.2356 & -0.1199 & 0.9644 \end{bmatrix} \begin{bmatrix} z_t \\ a_t \\ \widetilde{K}_t \end{bmatrix} + v_t,$$

where v_t is a two-dimensional vector of independent Gaussian noises with variances 0.0004 and 0.1347. As λ becomes smaller, acquiring information is less costly and hence the solution under RI is closer to that under full information.

As is well known in control theory, the optimal solution under full information does not depend on a particular choice of states and controls. For example, one can choose the control vector as $u_t = (I_t, N_t, e_t)'$ instead of $(C_t, N_t, e_t)'$. Then one can eliminate C_t using equation (30) and replace the state transition equation (34) for capital with equation (31). This procedure does

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

not affect the solution under full information but matters under RI. The intuition is that control variables must be known to the DM; that is, they must be adapted to the DM's information set $\{s^t\}$. This may force some other variables not adapted to the information set by the model constraints. For example, if one chooses $(I_t, N_t, e_t)'$ as the control vector, then consumption C_t may not be adapted to s^t because the resource constraint (30) must be satisfied. If both C_t and I_t were adapted to s^t, then output $\exp(a_t)\,(e_t K_t)^{\alpha} N_t^{1-\alpha}$ would be adapted too. But this is generally impossible because both a_t and K_t are unobserved states.

5.3. Planner's problem II: investment shock

In this subsection, we remove the preference shock in the previous example but introduce an investment shock. Under full information, the planner's problem is to maximize the following objective:

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t U(C_t, N_t)\right],\ U(C,N) = \log(C) - \chi\frac{N^{1+\nu}}{1+\nu},$$

subject to equation (30) and

$$\begin{aligned} K_{t+1} &= (1-\delta(e_t))\,K_t + \exp(z_t)\,I_t,\ (\Lambda_{3t}),\\ a_{t+1} &= \rho_a a_t + \epsilon_{a,t+1},\ (\Lambda_{2t}),\\ z_{t+1} &= \rho_z z_t + \epsilon_{z,t+1},\ (\Lambda_{1t}), \end{aligned} \tag{36}$$

where z_t represents an investment shock and Λ_{it}, $i = 1, 2, 3$ are the Lagrange multipliers associated with the state transition equations for z_t, a_t, and K_t. Here, $\epsilon_{a,t}$ and $\epsilon_{z,t}$ are independent Gaussian white noise with variances σ_a^2 and σ_z^2.

We choose $x_t = (z_t, a_t, K_t)'$ as the state vector and $u_t = (C_t, N_t, e_t)'$ as the control vector. Use equation (30) to substitute for I_t in equation (36). The steady state values of Λ_{it}, $i = 1, 2, 3$, satisfy

$$\overline{\Lambda}_1 = \frac{\beta\bar{I}}{(1-\beta\rho_z)\,\overline{C}},\ \overline{\Lambda}_2 = \frac{\beta U_c\left(\overline{C},\overline{N}\right)\overline{K}^{\alpha}\overline{N}^{1-\alpha}}{1-\beta\rho_a},\ \overline{\Lambda}_3 = U_c\left(\overline{C},\overline{N}\right),$$

given the steady state capital utilization rate $\bar{e} = 1$.

We set the parameter values as $\alpha = 0.33$, $\delta = 0.025$, $\nu = 1$, $\beta = 0.99$, $\gamma = 1.2$, $\rho_a = 0.95$, $\rho_z = 0.7$, $\sigma_a = 0.01$, $\sigma_z = 0.01$, and $\lambda = 0.002$. We choose $\chi = 7.8827$ and $\phi_e = 0.0351$ such that the steady state labor and capital utilization rate are equal to $1/3$ and 1, respectively. The coefficient matrix in the optimal linearized policy function $\widetilde{u}_t = -F\mathbb{E}\left[\widetilde{x}_t|s^t\right]$ is given by:

$$F = \begin{bmatrix} 0.6716 & -0.3520 & -0.0213\\ -0.3881 & -0.2855 & 0.0061\\ -2.0462 & -1.8091 & 0.0955 \end{bmatrix}.$$

The matrix W is singular, but A is invertible and thus $AA' + W$ is invertible. As Assumption 2 is satisfied, our toolbox can be applied to compute the optimal information structure under RI. We find the optimal posterior covariance matrix for the state is given by:

$$\Sigma = \begin{bmatrix} 0.0195 & -0.0018 & -0.0028\\ -0.0018 & 0.0335 & 0.0335\\ -0.0028 & 0.0335 & 1.7259 \end{bmatrix} \times 10^{-2}.$$

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

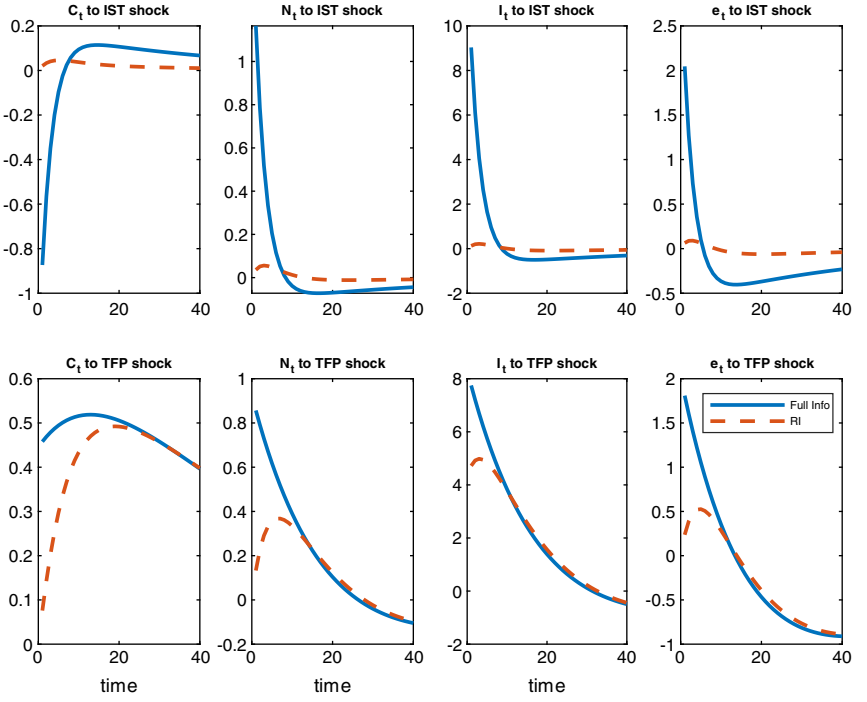


Figure 3. Impulse responses of I_t, N_t, and C_t to a one-percent innovation in the TFP and investment shocks under full information and under RI. All vertical axes are measured in percentage changes from the steady state.

and the optimal signal takes the following form:

$$s_t = 0.2594 z_t + 0.9632 a_t + 0.0706 \widetilde{K}_t + \nu_t, \tag{37}$$

where ν_t is a Gaussian white noise with variance 0.0022. The signal is one-dimensional and indicates that the planner may attribute a positive investment shock to a positive TFP shock given a positive realization of the signal.

Figure 3 presents the IRFs to the two shocks. As is well known, the investment shock cannot generate comovement between consumption and investment under full information. In this case, a positive investment shock causes investment to rise, but consumption to fall. By contrast, under RI, the planner confuses an investment shock with a TFP shock given the optimally chosen signal of the form (37). Thus, the investment responses to either a positive TFP shock or a positive investment shock are damped and delayed. Moreover, consumption rises on impact following a positive investment shock, generating comovement with investment.

Again endogenous capital utilization plays an important role. With exogenous capital utilization with $e_t = 1$ for all t, the model under RI could not generate comovement. In response to a positive investment shock, both investment and capital utilization rise, causing output to rise more beyond the rise of labor. This allows consumption to rise.

5.4. Planner's problem III: news shock

We now remove the investment shock in the previous example and consider a news shock. Under full information, the planner's problem is to maximize the following objective:

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t U(C_t, N_t)\right], \; U(C, N) = \log(C) - \chi \frac{N^{1+\nu}}{1+\nu},$$

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

subject to equations (30), (31), and

$$a_{t+1} = \rho_a a_t + \epsilon_{a,t+1} + \epsilon_{n,t-h},$$

where $\epsilon_{n,t-h}$ is a Gaussian white noise with variance σ_n^2. The noise $\epsilon_{n,t}$ represents a news shock to the future TFP that is announced at date t but realized in $h+1$ period later.

This model does not directly fit into our general framework. We use a simple example with $h=2$ to illustrate how this model can be transformed into our framework. We define the state vector as $x_t = \left(a_t, K_t, y_{1t}, y_{2t}, y_{3t}\right)'$ and the control vector as $u_t = (C_t, N_t, e_t)'$. The state transition equations become

$$\begin{aligned}
a_{t+1} &= \rho_a a_t + \epsilon_{a,t+1} + y_{1t}, \quad (\Lambda_{1t}) \\
K_{t+1} &= (1-\delta(e_t))\, K_t + \exp(a_t)\,(e_t K_t)^{\alpha} N_t^{1-\alpha} - C_t, \, (\Lambda_{2t}) \\
y_{1,t+1} &= y_{2t}, \quad (\Lambda_{3t}) \\
y_{2,t+1} &= y_{3t}, \quad (\Lambda_{4t}) \\
y_{3,t+1} &= \epsilon_{n,t+1}, \quad (\Lambda_{5t})
\end{aligned}$$

where Λ_{it}, $i = 1, 2, .., 5$, represent the Lagrange multipliers associated with these equations. Only the steady state value of Λ_{2t} matters for the LQ approximation as the others are associated with linear constraints.

We choose the same parameter values as in Section 5.1 except for $\lambda = 0.005$. In addition, we set $\sigma_n = 1\%$. Using our toolbox, we derive the coefficient matrix in the linearized policy function:

$$F = \begin{bmatrix} -0.3520 & -0.0213 & -0.3293 & -0.3184 & -0.3078 \\ -0.2855 & 0.0061 & 0.1327 & 0.1283 & 0.1240 \\ -1.8091 & 0.0955 & 0.3066 & 0.2964 & 0.2865 \end{bmatrix}.$$

The optimal signal takes the form:

$$s_t = 0.6619 a_t + 0.0490 K_t + 0.5340 \epsilon_{n,t-2} + 0.4184 \epsilon_{n,t-1} + 0.3151 \epsilon_{n,t} + \nu_t, \tag{38}$$

where ν_t is a Gaussian white noise with variance 0.0026. This signal form is similar to that in Maćkowiak et al. (2018). Their model does not have the endogenous capital state. Maćkowiak and Wiederholt (2020) include capital in an equilibrium model and show that the household's or firm's decision problem can be approximated by an LQ tracking problem. These two papers show that RI can help explain the comovement puzzle. Unlike these papers, here we analyze a social planner's control problem in which K_t is a hidden state and is included in the signal.

As is well known, news shock cannot generate comovement in a standard RBC model under full information [Beaudry and Portier (2004)]. As shown in Figure 4, a positive news shock about future TFP raises consumption but reduces labor and investment. Introducing variable capital utilization alone does not help because the capital utilization declines making output declines even more. Thus, a positive news shock cannot cause both consumption and investment to rise. To generate comovement in the RBC framework, Jaimovich and Rebelo (2009) argue that one has to introduce two more elements: preferences that allow the modeler to parameterize the strength of wealth effects and investment adjustment costs.

By contrast, Figure 4 shows that RI combined with variable capital utilization can generate comovement. The intuition comes from the optimal signal form in equation (38), which implies that the planner cannot distinguish a positive news shock about the future TFP from a positive shock to the current TFP. Thus, the planner raises labor supply, investment, and capital utilization in response to a positive news shock. The model also generates damped, delayed, and hump-shaped responses to the current TFP shock, even though there is no investment adjustment cost.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

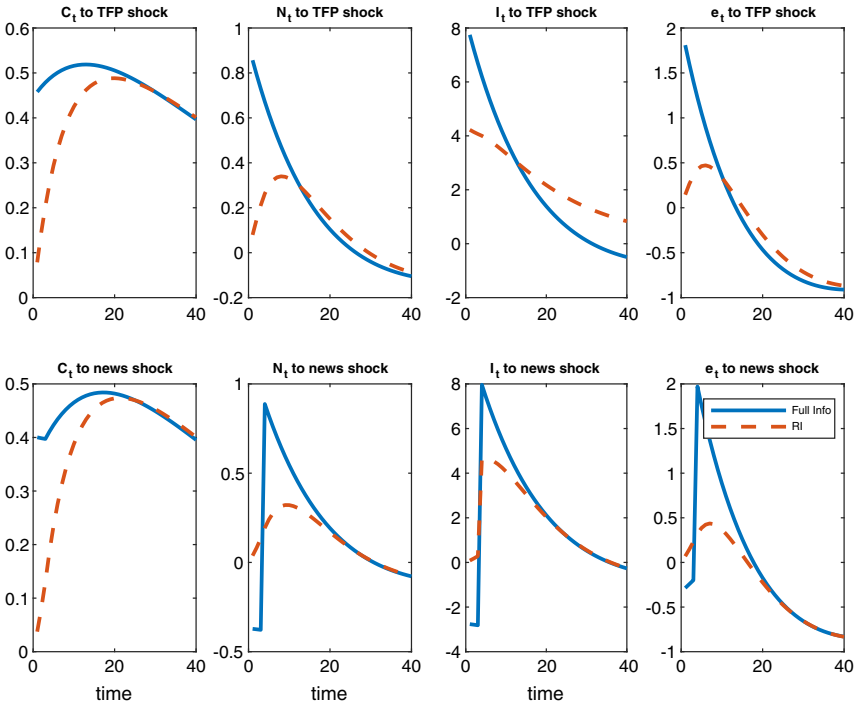


Figure 4. Impulse responses of I_t, N_t, and C_t to a one-percent innovation in the TFP and news shocks under full information and under RI. All vertical axes are measured in percentage changes from the steady state.

5.5. Durable and nondurable consumption

In this subsection, we study an agent's consumption/saving problem with both durable and nondurable goods. We modify the models of Bernanke (1985) and Luo et al. (2015) by introducing a general power utility function and adjustment costs.

Under full information, the agent chooses nondurable consumption $\{c_t\}$, durable good investment $\{I_t\}$, and asset holdings $\{b_{t+1}\}$ to maximize discounted utility:

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[U(c_t,k_t)-\frac{\phi_b}{2}\left(b_t-\bar{b}\right)^2\right],\ U(c,k)=\frac{\left(c^{\theta}k^{1-\theta}\right)^{1-\gamma}}{1-\gamma},$$

subject to

$$\begin{aligned} c_t+b_{t+1}+I_t+\frac{\phi_k k_t}{2}\left(\frac{I_t}{k_t}-\delta\right)^2 &= Rb_t+\bar{y}\exp(y_{1t}+y_{2t}+\epsilon_{zt}),\ \Lambda_{5t}\\ k_{t+1} &= (1-\delta)\,k_t+I_t,\ (\Lambda_{4t})\\ y_{1,t+1} &= \rho_1 y_{1t}+\epsilon_{1,t+1},\ (\Lambda_{1,t})\\ y_{2,t+1} &= \rho_2 y_{2t}+\epsilon_{2,t+1},\ (\Lambda_{2t}) \end{aligned} \tag{39a}$$

where $\beta=1/R$, k_t represents durable consumption, and ϵ_{1t}, ϵ_{2t}, and ϵ_{zt} are independent Gaussian white noise with variances σ_1^2, σ_2^2, and σ_z^2, respectively. Here, y_{1t} and y_{2t} are two persistent components of labor income and ϵ_{zt} is a purely temporary component.

Assume that durable investment incurs quadratic costs with parameter $\phi_k>0$. Following Schmitt-Grohé and Uribe (2003), we also introduce portfolio adjustment costs with parameter $\phi_b>0$ to ensure that there is a nonstochastic steady state for the model given $\beta=1/R$. One

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

can show that the steady state asset holdings are $\bar{b}$ and the steady state nondurable and durable consumption is given by:

$$\bar{c} = \frac{(R-1)\bar{b} + \bar{y}}{1 + \frac{\delta\beta(1-\theta)}{\theta[1-\beta(1-\delta)]}} > 0,\ \bar{k} = \frac{\beta(1-\theta)}{\theta\left[1-\beta(1-\delta)\right]}\bar{c}.$$

This is a version of the permanent income hypothesis in that both durable and nondurable consumption levels are proportional to the annuity value of his human and nonhuman wealth $(R-1)\,\bar{b} + \bar{y}$ in the steady state.

When there is no adjustment cost ($\phi_k = \phi_b = 0$) and when utility is quadratic, for example, $U(c,k) = -(c - c_{\max})^2 - \theta(k - k_{\max})^2$, Luo et al. (2015) show that one can choose the expected lifetime resource (analogous to permanent income) as a single state variable so that optimal dural and nondurable consumption is linear in this state variable. As both consumption processes follow a random walk, there is no nonstochastic steady state. Unfortunately, their approach does not apply to models beyond the LQ framework, like our model.

We now apply our LQ approximation approach. We choose the state vector as $x_t = (y_{1t}, y_{2t}, y_{zt}, k_t, b_t)'$ and the control vector as $u_t = (c_t, I_t)'$. The state variable y_{zt} represents the temporary shock ϵ_{zt}, and its transition equation is given by:

$$y_{z,t+1} = \epsilon_{z,t+1}.\ (\Lambda_{3t})$$

Replace ϵ_{zt} with y_{zt} in (39a). Then the model fits into our general framework. Let Λ_{it}, $i = 1, \ldots, 5$, denote the Lagrange multipliers associated with the state transition equations for x_t. One can derive their steady state values as:

$$\bar{\Lambda}_1 = \frac{\beta\bar{y}}{1-\beta\rho_1}U_c,\ \bar{\Lambda}_2 = \frac{\beta\bar{y}}{1-\beta\rho_2}U_c,$$
$$\bar{\Lambda}_3 = \beta\bar{y}U_c,\ \bar{\Lambda}_4 = \bar{\Lambda}_5 = U_c,\ U_c = \left(\bar{c}^{\theta}\bar{k}^{1-\theta}\right)^{-\gamma}\theta\bar{c}^{\theta-1}\bar{k}^{1-\theta}.$$

Suppose that the agent does not fully observe the states and acquires information to learn about the states subject to discounted information costs modeled by the discounted mutual information. To solve the model numerically, we set the following parameter values: $R = 1.01$, $\beta = 1/R$, $\gamma = 2$, $\theta = 0.8$, $\bar{y} = 1$, $\bar{b} = 0.2$, $\delta = 0.02$, $\phi_k = 0.5$, $\phi_b = 0.001$, $\rho_1 = 0.97$, $\rho_2 = 0.8$, $\sigma_1^2 = 0.0001$, $\sigma_2^2 = 0.003$, $\sigma_z^2 = 0.01$, and $\lambda = 0.005$.

Our toolbox delivers the coefficient matrix in the linearized policy function:

$$F = \begin{bmatrix} -0.3980 & -0.1037 & -0.0234 & -0.0071 & -0.0236 \\ -0.6889 & -0.2660 & -0.0851 & 0.1489 & -0.0860 \end{bmatrix}.$$

The optimal signal is given by:

$$s_t = 0.9593\widetilde{y}_{1t} + 0.2678\widetilde{y}_{2t} + 0.0604\widetilde{y}_{zt} - 0.0240\widetilde{k}_t + 0.0610\widetilde{b}_t + v_t, \tag{40}$$

where v_t is a Gaussian white noise with variance 0.0129.

Figure 5 shows IRFs to a one-percent positive innovation shock to y_{1t}. Under full information, both nondurable consumption c_t and the durable good expenditure I_t rise on impact and decline monotonically to the steady state. By contrast, both responses under RI are damped, delayed, and have a hump shape. Figures 6 and 7 present IRFs to a one-percent positive innovation shock to the less persistent income y_{2t} and the purely temporary income y_{zt}. These figures show that they have similar patterns, though the magnitudes of responses are different. The intuition comes from the optimal signal form in (40). The agent cannot differentiate three sources of income shocks given this signal.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

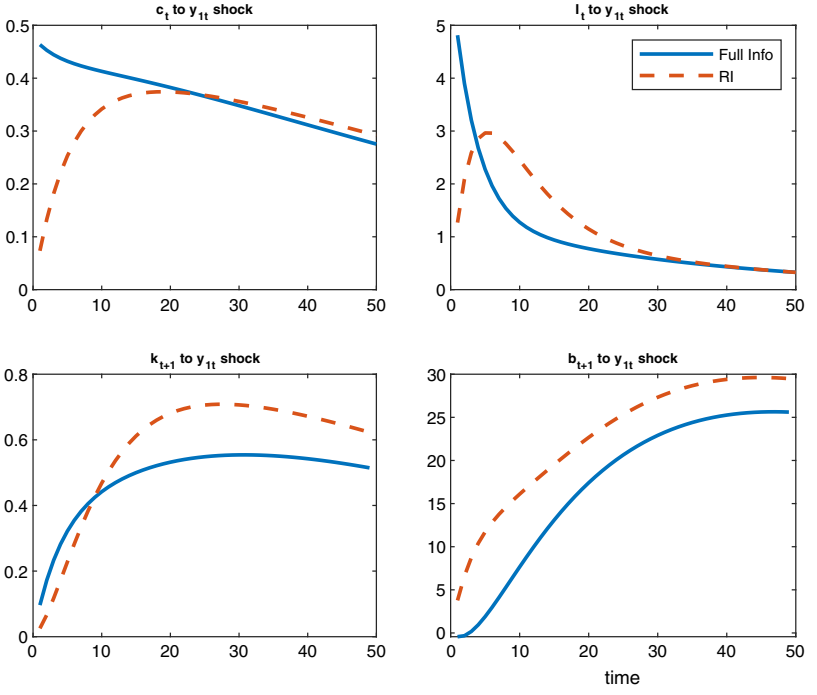


Figure 5. Impulse responses to a one-percent innovation shock to y_{1t} under full information and under RI. All vertical axes are measured in percentage changes from the steady state.

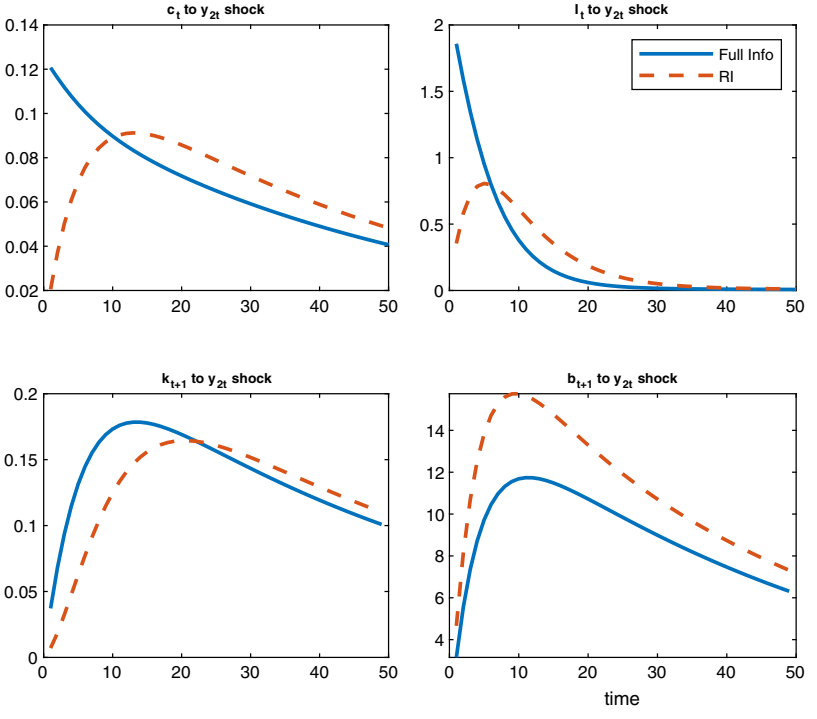


Figure 6. Impulse responses to a one-percent innovation shock to y_{2t} under full information and under RI. All vertical axes are measured in percentage changes from the steady state.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

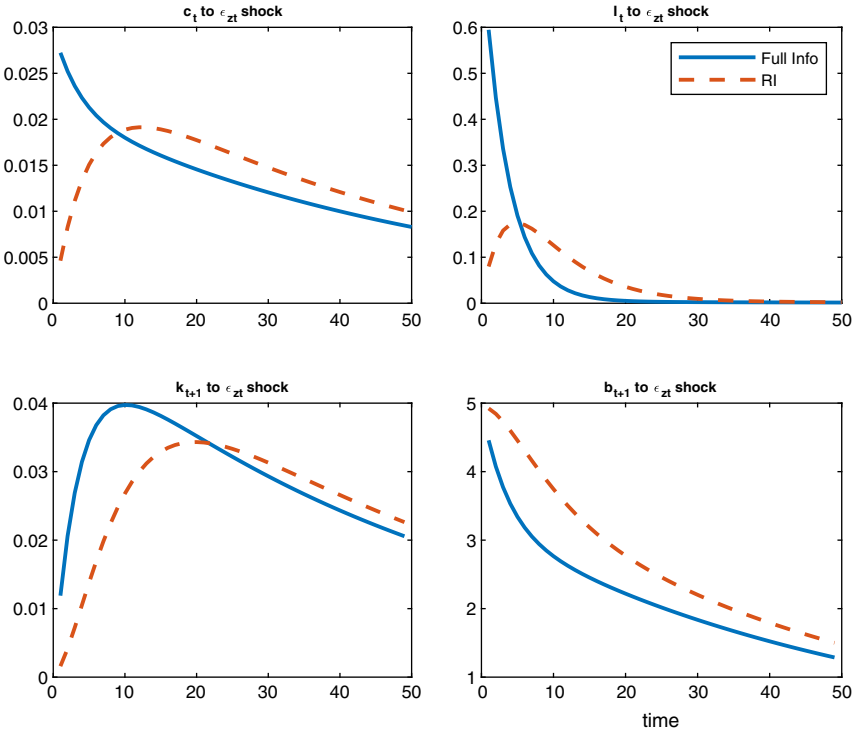


Figure 7. Impulse responses to a one-percent innovation shock to ϵ_{zt} under full information and under RI. All vertical axes are measured in percentage changes from the steady state.

6. Conclusion

In this paper, we have proposed an LQ approximation approach to dynamic nonlinear rationally inattentive control problems with multiple states and multiple controls. We have also provided a toolbox to implement this approach efficiently. Applying our toolbox to five economic examples, we show that RI possibly combined with variable capital utilization can help explain the comovement puzzle in the DSGE literature.

While our approach is quite general and can be easily implemented just as in the DSGE literature, some qualifications and weaknesses need to be stressed. First, our approach is local in nature and needs the problem to be smooth and the solution to be interior. Thus, our approach does not apply to problems with occasionally binding constraints or nonsmooth objective functions. Second, our approach applies to decision problems under Gaussian uncertainty only. Solving market equilibrium problems needs an outer loop to find equilibrium prices. Third, our approach obtains a correct linearized solution around a deterministic steady state. As Kim and Kim (2003) point out, when this linearized solution is used to evaluate welfare, it can lead to spurious results. Finally, because the certainty equivalence principle holds in our approach, the equilibrium distribution from the linearized solution does not take into account the feedback of risk on the DMs' behavior in the nonlinear control problems.

Notes

1 See Sims (2011), Angeletos and Lian (2017), and Maćkowiak et al. (2020) for surveys.

2 There is also a literature that studies rationally inattentive discrete choice problems [e.g. Woodford (2009), Matějka and McKay (2015), Steiner et al. (2017), Caplin et al. (2019), and Miao and Xing (in press)]. Sims (2006), Mondria (2010), Kacperczyk et al. (2016), Kőszegi and Matějka (2020), Hébert and La'O (2021), Fulton (2022), and Miao and Su (2023) study

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

static RI models with continuous choices. By contrast, our paper focuses on dynamic control problems with continuous choices.

3 See, for example, Zorn (2018) and Maćkowiak and Wiederholt (2015) for applications of this approach.

4 It is defined as:

$$H(X|Y) = \int \int p(x, y) \log p(x|y) dxdy,$$

where $p(x, y)$ and $p(x|y)$ are the joint pdf of (X, Y) and the conditional pdf of X given Y. See Cover and Thomas (2006).

5 There is also a transversality condition:

$$\lim_{t \to \infty} \beta^t \mathbb{E} [\lambda_t x_{t+1}] = 0.$$

6 We use the conventional matrix inequality notations: $X \succ (\succeq) Y$ means that $X - Y$ is positive definite (semidefinite) and $X \prec (\preceq) Y$ means $X - Y$ is negative definite (semidefinite).

7 See Angeletos and Lian (2022) for a discussion of the importance of both partial information and variable capital utilization for generating comovement.

References

Afrouzi, H. and C. Yang (2021) Dynamic Rational Inattention and the Phillips Curve. Working Paper, Columbia University.

Angeletos, G.-M. and C. Lian (2017) Incomplete information in macroeconomics: Accommodating frictions in oordination. In: Angeletos, G.-M. and C. Lian (eds.), *Handbook of Macroeconomics*, vol. 2, pp. 1065–1240. Amsterdam: North Holland.

Angeletos, G.-M. and C. Lian (2022) Confidence and the propagation of demand shocks. *Review of Economic Studies* 89(3), 1085–1119.

Beaudry, P. and F. Portier (2004) An exploration into Pigou's theory of cycles. *Journal of Monetary Economics* 51(6), 1183–1216.

Benigno, P. and M. Woodford (2012) Linear-quadratic approximation of optimal policy problems. *Journal of Economic Theory* 147(1), 1–42.

Bernanke, B. (1985) Adjustment costs, durable, and aggregate consumption. *Journal of Monetary Economics* 15(1), 41–68.

Caplin, A., M. Dean and J. Leahy (2019) Rational inattention, optimal consideration sets, and stochastic choice. *Review of Economic Studies* 86(3), 1061–1094.

Cover, T. M. and J. A. Thomas (2006) Elements of Information Theory, 2nd ed. New York: John Wiley & Sons, Inc.

Fulton, C. (2022) Choosing what to pay attention to. *Theoretical Economics* 17(1), 153–184.

Hébert, B. and J. La'O (2021) Information Acquisition, Efficiency, and Non-fundamental Volatility. Working Paper, Columbia University and Stanford University.

Jaimovich, N. and S. Rebelo (2009) Can news about the future drive the business cycle? *American Economic Review* 99(4), 1097–1118.

Judd, K. L. (1998) *Numerical Methods in Economics*. Cambridge, MA: The MIT Press.

Kacperczyk, M., S. Van Nieuwerburgh and L. Veldkamp (2016) A rational theory of mutual funds' attention allocation. *Econometrica* 84(2), 571–626.

Kim, J. and S. Kim (2003) Spurious welfare reversals in international business cycle models. *Journal of International Economics* 60(2), 471–500.

Kydland, F. E. and E. C. Prescott (1982) Time to build and aggregate fluctuations. *Econometrica* 50(6), 1345–1370.

Kőszegi, B. and F. Matějka (2020) Choice simplification: A theory of mental budgeting and naive diversification. *Quarterly Journal of Economics* 135(2), 1153–1207.

Levine, P., J. Pearlman and R. Pierse (2008) Linear-quadratic approximation, external habit, and targeting rules. *Journal of Economic Dynamics and Control* 32(10), 3315–3349.

Luo, Y. (2008) Consumption dynamics under information processing constraints. *Review of Economic Dynamics* 11(2), 366–385.

Luo, Y., J. Nie and E. R. Young (2015) Slow information diffusion and the inertial behavior of durable consumption. *Journal of the European Economic Association* 13(5), 805–840.

Magill, M. J. P. (1977) Some new results on the local stability of the process of capital accumulation. *Journal of Economic Theory* 15(1), 174–210.

Matějka, F. and A. McKay (2015) Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review* 105(1), 272–298.

Maćkowiak, B., F. Matějka and M. Wiederholt (2018) Dynamic rational inattention: Analytical results. *Journal of Economic Theory* 176, 650–692.

Maćkowiak, B., F. Matějka and M. Wiederholt (2020) Rational Inattention: A Review. Working Paper, Sciences Po., CEPR DP1540.

Maćkowiak, B. and M. Wiederholt (2009) Optimal sticky prices under rational inattention. *American Economic Review* 99(3), 769–803.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

Maćkowiak, B. and M. Wiederholt (2015) Business cycle dynamics under rational inattention. *Review of Economic Studies* 82(4), 1502–1532.
Maćkowiak, B. and M. Wiederholt (2020) Rational Inattention and the Business Cycle Effects of Productivity and News Shocks. Working Paper, European Central Bank and Sciences Po.
Miao, J. and D. Su (2023) Asset market equilibrium under rational inattention. *Economc Theory* 75(1), 1–30.
Miao, J. and J. Wu (2021) A Matlab Toolbox to Solve Dynamic Multivariate Rational Inattention Problems in the LQG Framework. Working Paper, Boston University.
Miao, J., J. Wu and E. R. Young (2022) Multivariate rational inattention. *Econometrica* 2(2), 907–945.
Miao J. and H. Xing (in press) Dynamic discrete choice under rational inattention. *Economic Theory*.
Mondria, J. (2010) Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory* 145(5), 1837–1864.
Peng, L. (2005) Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis* 40(2), 307–329.
Peng, L. and W. Xiong (2006) Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80(3), 563–602.
Schmitt-Grohé, S. and M. Uribe (2003) Closing small open economy models. *Journal of Monetary Economics* 61, 163–185.
Sims, C. A. (1998). Stickiness. Carnegie-Rochester Conference Series on Public Policy 49, 317–356.
Sims, C. A. (2003) Implications of rational inattention. *Journal of Monetary Economics* 50(3), 665–690.
Sims, C. A. (2006) Rational inattention: Beyond the linear-quadratic case. *American Economic Review* 96(2), 158–163.
Sims, C. A. (2011) Rational inattention and monetary economics. In: B. M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics*, vol. 3A. Amsterdam: North-Holland.
Steiner, J., C. Stewart and F. Matějka (2017) Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica* 85(2), 521–553.
Tanaka, T., K. K. Kim, P. A. Parrio and S. K. Mitter (2017) Semidefinite programming approach to Gaussian sequential rate-distortion trade-offs. *IEEE Transactions on Automatic Control* 62(4), 1896–1910.
Woodford, M. (2009) Information-constrained state-dependent pricing. *Journal of Monetary Economics* 56, S100–S124.
Zorn, P. (2018) Investment Under Rational Inattention: Evidence from US Sectoral Data. Working Paper, University of Munich.

Appendix A: Proofs

Proof of Lemma 1: Define the function H as in (13). Using (1), we derive

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^t f(x_t,u_t)=\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[f(x_t,u_t)-\overline{\Lambda}'\left(x_{t+1}-g(x_t,u_t,\epsilon_{t+1})\right)\right].$$

Take a second-order Taylor expansion around the nonstochastic steady state. The constant term is

$$\frac{f(\bar{x},\bar{u})}{1-\beta}.$$

The first-order terms are equal to

$$\begin{aligned}&\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[f_x\widetilde{x}_t+f_u\widetilde{u}_t-\overline{\Lambda}'\left(\widetilde{x}_{t+1}-g_x\widetilde{x}_t-g_u\widetilde{u}_t-g_\epsilon\widetilde{\epsilon}_{t+1}\right)\right]\\&=\mathbb{E}\sum_{t=0}^{\infty}\beta^t\widetilde{u}_t\left(f_u+\overline{\Lambda}'g_u\right)+\mathbb{E}\left(f_x+\overline{\Lambda}'g_x\right)\widetilde{x}_0+\mathbb{E}\left[\beta\left(f_x+\overline{\Lambda}'g_x\right)\widetilde{x}_1-\overline{\Lambda}'\widetilde{x}_1\right]\\&\quad+\beta\mathbb{E}\left[\beta\left(f_x+\overline{\Lambda}'g_x\right)\widetilde{x}_2-\overline{\Lambda}'\widetilde{x}_2\right]+\ldots+\lim_{T\to\infty}\overline{\Lambda}'\beta^T\mathbb{E}\left[\widetilde{x}_{T+1}\right].\end{aligned}$$

Clearly, $\mathbb{E}\widetilde{x}_0=0$ as x_0 is exogenously given and $\mathbb{E}\widetilde{\epsilon}_{t+1}=0$ as $\widetilde{\epsilon}_{t+1}=\epsilon_{t+1}$. It follows from (7), (8), and $\lim_{T\to\infty}\mathbb{E}\left[\widetilde{x}_{T+1}\right]=0$ that the above first-order terms are equal to zero.

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press

The second-order terms are equal to

$$\frac{1}{2}\mathbb{E}\sum_{t=0}^{\infty}\beta^t\begin{bmatrix}\widetilde{x}_t\\ \widetilde{u}_t\\ \widetilde{\epsilon}_{t+1}\end{bmatrix}'\begin{bmatrix}H_{xx} & H_{xu} & H_{x\epsilon}\\ H_{ux} & H_{uu} & H_{u\epsilon}\\ H_{\epsilon x} & H_{ux} & H_{\epsilon\epsilon}\end{bmatrix}\begin{bmatrix}\widetilde{x}_t\\ \widetilde{u}_t\\ \widetilde{\epsilon}_{t+1}\end{bmatrix}$$

$$=\frac{1}{2}\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\widetilde{x}_t',\widetilde{u}_t'\right]\begin{bmatrix}H_{xx} & H_{xu}\\ H_{ux} & H_{uu}\end{bmatrix}\begin{bmatrix}\widetilde{x}_t\\ \widetilde{u}_t\end{bmatrix}+\frac{1}{2}\mathrm{Tr}(H_{\epsilon\epsilon}),$$

because $\mathbb{E}\epsilon_{t+1}=0$ and

$$\mathbb{E}\left[\widetilde{\epsilon}_{t+1}'H_{\epsilon\epsilon}\widetilde{\epsilon}_{t+1}\right]=\mathrm{Tr}\left(H_{\epsilon\epsilon}\mathbb{E}\left[\widetilde{\epsilon}_{t+1}\widetilde{\epsilon}_{t+1}'\right]\right)=\mathrm{Tr}(H_{\epsilon\epsilon}).$$

Combining the terms above yields the desired result. Q.E.D. □

Proof of Lemma 2: Define the Lagrangian as:

$$\frac{1}{2}\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\widetilde{x}_t',\widetilde{u}_t'\right]\begin{bmatrix}H_{xx} & H_{xu}\\ H_{ux} & H_{uu}\end{bmatrix}\begin{bmatrix}\widetilde{x}_t\\ \widetilde{u}_t\end{bmatrix}-\mathbb{E}\sum_{t=0}^{\infty}\beta^t\Phi_t'\left[\widetilde{x}_{t+1}-g_x\widetilde{x}_t-g_u\widetilde{u}_t-g_\epsilon\epsilon_{t+1}\right],$$

where Φ_t denotes the Lagrange multiplier adapted to s^t. The first-order conditions are given by:

$$\widetilde{u}_t:\ H_{uu}\widetilde{u}_t+H_{ux}\mathbb{E}_t\left[\widetilde{x}_t\right]+g_u'\Phi_t=0,$$

$$\widetilde{x}_{t+1}:\ \Phi_t=\beta\mathbb{E}_t\left[H_{xx}\widetilde{x}_{t+1}+H_{xu}\widetilde{u}_{t+1}+g_x'\Phi_{t+1}\right].$$

By the definition of H in equation (13), we deduce that the above conditions are the same as equations (9) and (10) by setting $\Phi_t=\widetilde{\Lambda}_t$. Q.E.D. □

Cite this article: Miao J and Zhang B (2024). "Linear quadratic approximation of rationally inattentive control problems." *Macroeconomic Dynamics* **28**, 1371–1393. https://doi.org/10.1017/S1365100523000500

https://doi.org/10.1017/S1365100523000500 Published online by Cambridge University Press