Monetary Policy with Near-Rational Expectations in Open Economies*

Yang Jiao† Seunghoon Na§ Anurag Singh¶

This version: March 28, 2022

Abstract

We investigate robustly optimal monetary policy in a small open economy New Keynesian model where private agents form potentially distorted near-rational expectations (NRE). Following a cost-push shock, the optimal monetary policy calls for sluggish domestic inflation and higher initial devaluation rate responses as the central bank's concerns about NRE increase. We estimate the models for Canada and Mexico. Mexico exhibits significant deviations from the rational expectation (RE) benchmark, whereas Canada shows small deviations. The NRE model for Mexico successfully predicts the dynamics of the monetary policy rate and moments of the CPI inflation rate, whereas the RE model fails.

JEL Classification: E52, E71, F41, F47.

Keywords. Expectation Formations, Heterogeneous Beliefs, Ambiguity, Monetary Policy, Small Open Economy, Bayesian Inference.

*We thank Yongsung Chang, Xiang Fang, Wataru Miyamoto, Woong Yong Park, Stephanie Schmitt-Grohé, Chang Sun, Cathy Zhang, Xiadong Zhu, and seminar participants at Purdue University, Seoul National University, and University of Hong Kong for their valuable comments. All errors are ours.

†Fanhai International School of Finance, Fudan University. Email: yang.jiao.0606@gmail.com.

§Department of Economics, Krannert School of Management, Purdue University. Email: na28@purdue.edu.

¶ITAM Business School, Email: anurag.singh@itam.mx

1 Introduction

The design of monetary policy in a globalized world has been a question of first-order importance. The vast open economy New Keynesian (NK) literature has typically conducted analyses based on the assumption of rational expectations (RE). However, there is growing consensus on widespread bounded rationality in economic agents' decision-making. In particular, various survey data and laboratory experiments find that forward-looking expectation formations of private sectors show substantial deviations from the RE. Then, how monetary policy should behave according to non-RE is a question of interest that requires a theoretical structure beyond the RE benchmark.

Investigating the macroeconomic consequences of agents without the RE has become an important research area in macroeconomic theory. One of the critical developments in the research area is introduced in Hansen and Sargent (2001) and Hansen and Sargent (2008), which formulates the departure of the model from the RE benchmark as an unknown probabilistic feature of uncertainty, motivated by the concept of *ambiguity* in Gilboa and Schmeidler (1989). Under the potential threat of a malevolent probabilistic environment under uncertainty about the model, decision-makers employ a *robust control* – the minimax optimization procedure – to minimize the worst-case outcome. Woodford (2010) and Adam and Woodford (2012) develop the idea further to design robustly optimal monetary policy in a standard closed-economy NK model, but private agents exhibit near-rational expectations (NRE). The NRE is potentially distorted from the RE with a given radius of relative entropy.

In this paper, we study the robustly optimal monetary policy in an open economy context. We start from the standard small open economy NK model as in Gali and Monacelli (2005). There are two types of goods in the economy: domestically produced goods and imported goods. Domestic producers use domestic labor as input and have the ability to set prices only in the domestic currency. The prices are not readjusted in an NK fashion. There is a trade of assets across countries in an international financial market. The model environment is summarized as equations for domestic aggregate demand and supply and equations for international prices and quantities. We extend the model environment as follows. We assume that the private agents exhibit NRE, and the central bank is the planner, who aims to maximize the private agents' welfare in a paternalistic fashion, but with a concern about the NRE. In addition, we assume the aggregate supply equation is shifted by cost-push shock, which creates non-trivial policy trade-offs between stabilizing inflation and the output gap.

When we solve the problem of robustly optimal monetary policy, we find that as the central bank has more concerns about the NRE, the robustly optimal monetary policy becomes more history-dependent in stabilizing inflation of domestically produced goods. When a cost-push shock hits the economy, the robustly optimal monetary policy demands a more sluggish reaction of domestic inflation as the central bank's concerns about the NRE in-

crease. In contrast, there is a more aggressive initial response of the nominal devaluation rate as a result of the optimal trade-offs between domestic inflation and the output gap. In addition, the optimal dynamics of the economy differ as the central bank's concerns about the NRE increase.

We estimate robustly optimal monetary policy models using time series data on CPI inflation rates and nominal devaluation rates for Canada and Mexico. We construct an algorithm to estimate the model via the likelihood-based Bayesian method. We find that there are significant differences between the estimated NRE models in the two countries. The NRE model shows minor deviations from the RE benchmark in Canada, whereas the NRE model in Mexico shows substantial deviations from its RE benchmark. More interestingly, the NRE model in Mexico successfully predicts the actual path of the monetary policy rate in the data, while the RE model fails. The central insight from the result is that the NRE model for Mexico successfully matches the key moments of the CPI inflation rate, which combines domestic inflation and exchange rate pass-through. In the estimated NRE model, the substantial deviation from the RE model motivates the central bank to adopt a more history-dependent monetary policy. The policy generates a large devaluation driven by the cost-push shock during the Mexican Peso Crisis, which results in high persistence in the dynamics of the CPI inflation. On the other hand, the estimated RE model has a limited ability to generate large devaluation as a policy outcome. The RE model predicts that the nominal devaluation during the Peso Crisis is mainly driven by the sharp rise of the domestic TFP shock, which is at odds with the fact that the crisis was not associated with technological progress.

To the best of our knowledge, this paper is the first to study the robustly optimal monetary policy in an open economy where the policymaker deals with private agents' NRE.[1] There is a vast body of work investigating robust monetary policy in a closed-economy New Keynesian environment (see, for example, Walsh (2004), Leitemo and Soderstrom (2008b), Dennis (2010), Levine and Pearlman (2010), Gerke and Hammermann (2016), and many others). There are also several related works on open-economy environments (see, for example, Dennis et al. (2009) and Leitemo and Soderstrom (2008a)). However, in these works, the optimization problem is typically based on the policymaker's doubts of her own model. Instead, we follow Woodford (2010) and Adam and Woodford (2012) to consider the type of

[1] As discussed in Hansen and Sargent (2012), there are three types of ambiguity in the problem of a policymaker facing ambiguity. Type-1 ambiguity is the case where the policymaker doubts both the exogenous shock processes and private agents' expectations. Type-2 ambiguity is the case where the policymaker doubts the model of the exogenous shock processes, whereas private agents fully trust the model. Finally, type-3 ambiguity is the case where the policymaker fully trusts the model of the exogenous shock processes, but it does not have full confidence in private agents' expectations. The type of ambiguity in our paper is type-3. Hansen and Sargent (2012) and Kwon and Miao (2017) consider the robustly optimal monetary policy problems under type-3 ambiguity, but their applications are closed-economy cases.

ambiguity under which the policymaker trusts the model of the exogenous shock processes, but it does not have full confidence in private agents' expectations. There is a lack of knowledge about the features of the robustly optimal monetary policy in an open economy in this case, and this paper attempts to fill this gap.

Furthermore, this paper measures the deviation from the RE benchmark by estimating the degree of NRE. Several papers estimate the deviation in various environments of boundedly rational agents. Ilut and Schneider (2014) apply the Bayesian method to estimate the household's ambiguity aversion, whereas the monetary policy follows the standard Taylor rule. Bhandari et al. (2019) estimate private agents' time-varying subjective beliefs by using survey data. Gust et al. (2020) apply the Bayesian method to estimate the averaged planning horizon of a New Keynesian environment with agents of limited foresight. Note that the information and policy structure of our paper differs essentially from these papers. To the best of our knowledge, our paper is also the first attempt to estimate the degree of the deviation from RE when the policymaker has Woodford (2010)'s type of ambiguity[2]. The optimal monetary policy is assumed to have a conditionally linear form, which enables us to summarize the model as a system of linear difference equations. Thus, we can apply standard techniques to estimate linear state-space models. However, the estimation of the system of the equilibrium with robustly optimal policy is computationally demanding, since the policy coefficients need to be computed for every draw in the Markov Chain in the estimation procedure. We construct and execute the estimation algorithm.

The remainder of the paper is organized as follows. Section 2 describes a small open economy model under near-rational private agents' expectations and robustly optimal monetary policy with commitment. Section 3 investigates the normative features of the robustly optimal monetary policy. Section 4 estimates the models by using the Bayesian method and discusses the key results. Section 5 concludes the paper.

2 Model

In this section, we incorporate the distorted expectations of private agents into the canonical small open economy New Keynesian model of Gali and Monacelli (2005). As is common in the New Keynesian literature, we further introduce a cost-push shock into the model and consider robustly optimal monetary policy responses. We begin by describing the belief structure of the private agents in the model.

[2]In such model environments, forecast errors from households' and firms' survey data cannot be used to identify the central bank's concerns regarding the private agents' NRE in a model-consistent way.

2.1 Uncertainty and Beliefs

The uncertainty of the economy is defined as the set of vectors of an exogenous stochastic disturbance process $\{\boldsymbol{\varepsilon}^t\}_{t=0}^{\infty}$, where $\boldsymbol{\varepsilon}^t \equiv \{\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_t\}$ is the history of periodic stochastic disturbance $\boldsymbol{\varepsilon}_t, \forall t \geq 0$.

We assume the following information structure depicted in Woodford (2010) and Adam and Woodford (2012). There is the probability triple $(\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{P}})$, where $\boldsymbol{\Omega}$ is the sample space, $\boldsymbol{\mathcal{F}}$ is the σ-Field of the sample space, and $\boldsymbol{\mathcal{P}}$ is the probability measure, defines the probability space of the exogenous process. The benchmark expectation operator is denoted by $\mathbb{E}[\cdot]$, which is induced by measure $\boldsymbol{\mathcal{P}}$. This is the *rational* measure for the probability of states of the economy. The private agents' expectations, however, are not necessarily rational in the sense that their expectation operator is induced by a potentially different measure $\hat{\boldsymbol{\mathcal{P}}}$. We impose the restriction that the measure $\hat{\boldsymbol{\mathcal{P}}}$ is absolutely continuous with respect to the measure $\boldsymbol{\mathcal{P}}$. Let $p(\boldsymbol{\varepsilon})$ denote the unconditional probability density of $\boldsymbol{\varepsilon}$ from measure $\boldsymbol{\mathcal{P}}$, where $\boldsymbol{\varepsilon}$ is a dummy random vector with the same dimension as the number of entries of $\boldsymbol{\varepsilon}_t$. Let $\hat{p}\left(\boldsymbol{\varepsilon}|\boldsymbol{\varepsilon}^t\right)$ denote the one-step-ahead probability density for $\boldsymbol{\varepsilon}_{t+1}$, which is induced by measure $\hat{\boldsymbol{\mathcal{P}}}$, conditioned on date-$t$ information. The likelihood ratio between the two densities is

$$m_{t+1} = \frac{\hat{p}\left(\boldsymbol{\varepsilon}|\boldsymbol{\varepsilon}^t\right)}{p(\boldsymbol{\varepsilon})},$$

and m_{t+1} is nonnegative and

$$\mathbb{E}\left[m_{t+1}|\boldsymbol{\varepsilon}^t\right] = 1. \tag{1}$$

Following Hansen and Sargent (2008), we set $\mathcal{M}_0 = 1$ and recursively construct $\{\mathcal{M}_t\}$ such that

$$\mathcal{M}_{t+1} = m_{t+1}\mathcal{M}_t,$$

which implies

$$\mathcal{M}_{t+j} = \prod_{k=1}^{j} m_{t+k},$$

and it is a martingale process that satisfies

$$\mathbb{E}[\mathcal{M}_{t+j}|\boldsymbol{\varepsilon}^t] = \mathcal{M}_t.$$

The Radon-Nikodym theorem indicates that private agents' expectation $\hat{\mathbb{E}}[\cdot]$ of a random variable X_{t+j} induced by measure $\hat{\boldsymbol{\mathcal{P}}}$ conditioned on date-t information can be expressed as

the expectation induced by the measure $\boldsymbol{\mathcal{P}}$ for the augmented random variable X_{t+j}:

$$\hat{\mathbb{E}}[X_{t+j}|\boldsymbol{\varepsilon}^t] = \mathbb{E}\left[\frac{\mathcal{M}_{t+j}}{\mathcal{M}_t} X_{t+j} \middle| \boldsymbol{\varepsilon}^t\right],$$

where $\frac{\mathcal{M}_{t+j}}{\mathcal{M}_t}$ represents the Radon-Nikodym derivatives, which completely summarize the belief distortions. Then the one-step-ahead expectation of the random variable X_{t+1} induced by measure $\hat{\boldsymbol{\mathcal{P}}}$ is expressed as follows:

$$\hat{\mathbb{E}}[X_{t+1}|\boldsymbol{\varepsilon}^t] = \mathbb{E}[m_{t+1}X_{t+1}|\boldsymbol{\varepsilon}^t].$$

Henceforth, we simply express an expectation based on date-t information as the expectation with subscript t, i.e., $\hat{\mathbb{E}}_t X_{t+1} \equiv \hat{\mathbb{E}}[X_{t+1}|\boldsymbol{\varepsilon}^t]$ and $\mathbb{E}_t X_{t+1} \equiv \mathbb{E}[X_{t+1}|\boldsymbol{\varepsilon}^t]$.

We employ the relative entropy (Kullback-Leibler divergence) to measure the distance between two probability measures $\hat{\boldsymbol{\mathcal{P}}}$ and $\boldsymbol{\mathcal{P}}$. The distance of one-period-ahead distorted beliefs from the rational belief is summarized by the following relative entropy:

$$\mathcal{R}_t \equiv \mathbb{E}_t m_{t+1} \ln m_{t+1},$$

which is always nonnegative by Gibb's inequality.

In the next subsections, we use the operator $\hat{\mathbb{E}}[\cdot]$ to represent the private sector's distorted expectations.

2.2 A Small Open Economy NK Environment

Now, we consider a small open economy New Keynesian environment developed in Gali and Monacelli (2005) and assume that the agents' expectations are potentially distorted and that a cost-push shock exists. The aggregate demand of the economy can be represented by the following equation:

$$x_t = \hat{\mathbb{E}}_t x_{t+1} - \frac{1}{\sigma_\alpha}\left(i_t - \hat{\mathbb{E}}_t \pi_{H,t+1} - \bar{rr}_t\right), \tag{2}$$

where x_t is the domestic output gap defined by the actual output minus the natural level of output under flexible prices and $\pi_{H,t}$ is the net inflation rate for the domestically produced good. The variable i_t is the monetary policy rate determined by the policymaker. The parameter σ_α is defined as:

$$\sigma_\alpha \equiv \frac{\sigma}{1+\alpha(\omega-1)},$$

where $\sigma > 0$ governs the elasticity of intertemporal substitution in the private agents' utility functions, $\alpha \in [0, 1]$ governs the degree of home bias, and ω governs the effect of changes in the terms of trade on output. The variable $r\bar{r}_t$ is the natural rate of interest rate,

$$r\bar{r}_t = \rho + \Lambda_{r,a} a_t + \Lambda_{r,y^*} y_t^*, \tag{3}$$

where a_t denotes the domestic productivity shock and y_t^* is the world output shock. Parameter $\rho = -\log \beta$, and $\Lambda_{r,a}$ and Λ_{r,y^*} are coefficients that are functions of the structural parameters.

The aggregate supply of the economy is represented by the following equation:

$$\pi_{H,t} = \beta \hat{\mathbb{E}}_t \pi_{H,t+1} + \kappa x_t + u_t, \tag{4}$$

where $\beta \in (0, 1)$ refers to the subjective discount factor and u_t is the cost-push shock that shifts the aggregate supply curve. The parameter $\kappa > 0$ is defined as follows:

$$\kappa \equiv \frac{(1-\zeta)(1-\zeta\beta)(\sigma_\alpha + \varphi)}{\zeta},$$

where ζ is the Calvo-Yun parameter for nominal price rigidity and φ is the inverse of the Frisch-elasticity of labor supply.

We define the bilateral terms of trade $\mathcal{S}_{i,t} \equiv \frac{P_{i,t}}{P_{H,t}}$ as the ratio between the price of country i's good $P_{i,t}$ (in terms of domestic currency) and the price of domestic goods $P_{H,t}$. The effective terms of trade is given by $\mathcal{S}_t \equiv \frac{P_{F,t}}{P_{H,t}}$, where $P_{F,t}$ is the price index of the imported goods. The log of the effective terms of trade s_t is then expressed as $s_t = p_{F,t} - p_{H,t}$, where $p_{F,t}$ and $p_{H,t}$ are the logs of $P_{F,t}$ and $P_{H,t}$, respectively. Then the complete international asset market and international risk-sharing condition of Gali and Monacelli (2005) implies that the terms of trade s_t is expressed by the equation

$$s_t = \sigma_\alpha (x_t + \tilde{y}_t - y_t^*), \tag{5}$$

where $\tilde{y}_t$ is the natural level of domestic output,

$$\tilde{y}_t = \Lambda_{y,0} + \Lambda_{y,a} a_t + \Lambda_{y,y^*} y_t^*, \tag{6}$$

where $\Lambda_{y,0}, \Lambda_{y,a}$ and Λ_{y,y^*} are coefficients that are functions of the structural parameters (see Table C.6 in Appendix C for detailed description of $\Lambda_{y,0}, \Lambda_{r,a}, \Lambda_{r,y^*}, \Lambda_{y,a}$, and Λ_{y,y^*}). The

variable y_t^* is an exogenous world output shock.

We assume that the law of one price holds for individual goods, for both the import price and export price. The assumption results in the relation $P_{i,t} = \mathcal{E}_{it} P_{i,t}^i$, where $\mathcal{E}_{it}$ is the bilateral nominal exchange rate of the currency of country i in terms of domestic currency and $P_{i,t}^i$ is the price of country i's good in terms of country i's own currency. Then the log of the effective exchange rate between the domestic country and the rest of the world e_t is approximated by $e_t = p_{F,t} - p_t^*$, where p_t^* is the log of the effective price index of the rest of the world. By combining this with the definition of the log of the terms of trade, the nominal devaluation rate Δe_t is then expressed by the accounting equation

$$\Delta e_t = \pi_{H,t} + \Delta s_t - \pi_t^*, \tag{7}$$

where Δ is the first-order lagged difference operator and π_t^* is the exogenous world inflation rate. A positive (negative) Δe_t means a nominal depreciation (appreciation) of the domestic currency.

The three shocks described above are assumed to follow the following first-order Markov processes:

$$u_{t+1} = \rho_u u_t + \sigma_u \epsilon_{t+1}^u, \tag{8}$$

$$a_{t+1} = \rho_a a_t + \sigma_a \epsilon_{t+1}^a, \tag{9}$$

$$y_{t+1}^* = \rho_{y^*} y_t^* + \sigma_{y^*} \epsilon_{t+1}^y, \tag{10}$$

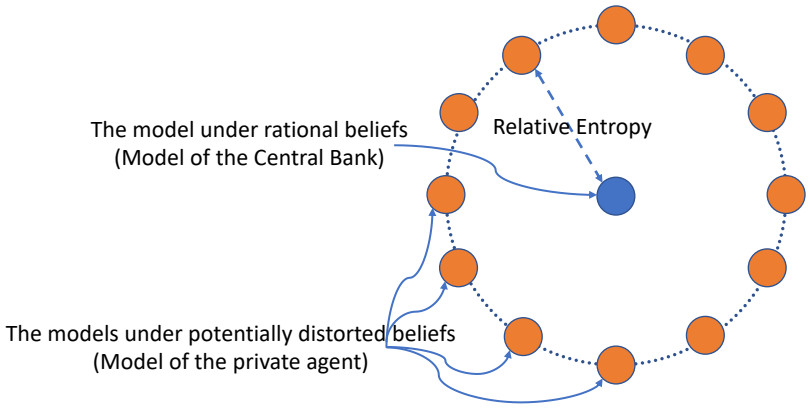
where parameter $\rho_i, i \in \{a, u, y^*\}$ governs the persistence, $\sigma_i, i \in \{a, u, y^*\}$ governs the standard deviation of each stochastic process, and $\varepsilon_{t+1}^i, i \in \{a, u, y^*\}$ follows an i.i.d standard normal distribution.

2.3 The Robustly Optimal Monetary Policy

Throughout this paper, we use the term 'concerns about distorted expectations' when we refer to the central bank's concerns regarding the potentially distorted expectations of private agents. As the planner of the economy, the central bank conducts optimal monetary policy by using the nominal interest rate i_t as the policy instrument. However, the optimal policy problem differs from the standard optimal policy problem under the RE since the belief structures of the central bank and private agents are heterogenous.

As in Woodford (2010), we focus on the case in which, unlike the private agents, the central bank has the rational belief. Figure 1 illustrates the environment that features a

Figure 1: Entropy Diagram



central bank and private agents with heterogeneous belief structures. In the figure, each ball refers to a model with distinct beliefs. The blue ball refers to the central bank's model, and the orange balls refer to the set of private agents' models. The set of private agents' models with potentially distorted expectations are around the central bank's model with RE, with the distance reflecting the relative entropy. The private agents have no concerns about their models when they make decisions. The central bank, to the contrary, is concerned about the private agents' potentially distorted expectations when it designs the optimal monetary policy. However, the central bank does not have further information on the private agents' expectations beyond the fact that they are potentially distorted. The decision-making environment under heterogeneous beliefs is consistent with the type-3 ambiguity aversion in Hansen and Sargent (2012).

Under the belief structure, the central bank seeks to minimize the private agents' welfare loss with a *paternalistic* objective (the welfare loss under rational belief) and with ambiguity aversion, which is introduced in Woodford (2010) and Adam and Woodford (2012). Specifically, the central bank's welfare loss function of the economy can be written as follows:

$$\mathbb{L} = \underbrace{\mathbb{E}_{-1}\sum_{t=0}^{\infty}\beta^t\frac{1}{2}\left(\pi_{H,t}^2+\lambda_x\left(x_t-\bar{x}\right)^2\right)}_{\text{the term of paternalistic welfare loss}} \underbrace{-\theta\mathbb{E}_{-1}\sum_{t=0}^{\infty}\beta^t m_{t+1}\ln m_{t+1}}_{\text{the term of ambiguity aversion}},$$

where the first term in parentheses on the right-hand side represents the discounted lifetime welfare loss of private agents under rational belief, which follows the welfare loss function in Gali and Monacelli (2005). The second term represents the concerns about distorted expectations. The parameter $\lambda_x > 0$ captures the weight of the welfare loss from the deviation of the output gap from its target value, $\bar{x}$, and $\theta \in (0, \infty)$ is related to the central bank's

concerns about distorted expectations. The welfare loss is a convex function with respect to $\pi_{H,t}, x_t$, but it is a concave function with respect to m_{t+1}.

Given its concerns about distorted expectations, the central bank attempts to minimize the welfare loss under the *worst-case* outcomes caused by the potentially distorted belief m_{t+1}. Following the framework of Woodford (2010), we assume that a hypothetically malevolent nature chooses $\{m_{t+1}\}$ to maximize the welfare loss, and given this, the policymaker chooses the best responses of $\{\pi_t, x_t, i_t\}$ to minimize the loss. We define the central bank's robustly optimal monetary policy problem as follows.

Definition 2.1 *The robustly optimal monetary policy of the central bank is solving the following minimax optimization problem:*

$$\min_{\{\pi_{H,t},x_t,i_t\}} \max_{\{m_{t+1}\}} \mathbb{E}_{-1} \sum_{t=0}^{\infty} \beta^t \frac{1}{2} \left(\pi_{H,t}^2 + \lambda_x \left(x_t - \bar{x}\right)^2\right) - \theta \mathbb{E}_{-1} \sum_{t=0}^{\infty} \beta^t m_{t+1} \ln m_{t+1}, \tag{11}$$

subject to equations (1), (2), (3), (4), (8), (9), and (10).

Note that when $\theta \to \infty$, the optimal behavior for the malevolent nature to maximize the welfare loss is to choose $m_{t+1} = 1$, which is the belief under the RE. When θ is a finite number, on the other hand, the optimal choice of m_{t+1} causes a departure from the belief under the RE. In this sense, higher $\theta^{-1} \in [0, \infty)$ can be interpreted as a higher degree of concerns about distorted expectations.

As a benevolent planner, the central bank commits to its past policy promises, which implies a history-dependent policy. The robustly optimal monetary policy problem defined in 2.1 consists of choosing a sequence of potentially time-varying $\{\pi_{H,t}, x_t\}_{t=0}^{\infty}$ which generally creates nonlinearity in the system. Instead, we employ the conditionally linear and self-consistent commitment policy developed in Woodford (2010) to maintain the linear system under commitment. Under conditionally linear commitment, solving the problem reduces to choosing a sequence of $\{\pi_{H,t}, x_t\}_{t=1}^{\infty}$ while taking the initial commitment $\{\pi_{H,0}, x_0\}$ as given. We suppose that the initial commitment takes a linear form:

$$\begin{bmatrix} \pi_{H,0} \\ x_0 \end{bmatrix} = \boldsymbol{\Phi}_{-1} + \boldsymbol{\Gamma}_{-1} \boldsymbol{\varepsilon}_0,$$

where $\boldsymbol{\Phi}_{-1} \equiv [\Phi_{\pi_H,-1}, \Phi_{x,-1}]'$ and $\boldsymbol{\Gamma}_{-1} \equiv [\Gamma_{\pi_H,u,-1}, \Gamma_{\pi_H,a,-1}, \Gamma_{\pi_H,y^*,-1}, \Gamma_{x,u,-1}, \Gamma_{x,a,-1}, \Gamma_{x,y^*,-1}]'$

complete an initial commitment. Then, we focus on the following conditionally linear rule:

$$\begin{bmatrix} \pi_{H,t+1} \\ x_{t+1} \end{bmatrix} = \boldsymbol{\Phi}_t + \boldsymbol{\Gamma}_t \boldsymbol{\varepsilon}_{t+1}, \ t \geq 0, \tag{12}$$

where $\boldsymbol{\Phi}_t \equiv [\Phi_{\pi_H,t}, \Phi_{x,t}]'$ is stochastic and $\boldsymbol{\Gamma}_t \equiv [\Gamma_{\pi_H,u,t}, \Gamma_{\pi_H,a,t}, \Gamma_{\pi_H,y^*,t}, \Gamma_{x,u,t}, \Gamma_{x,a,t}, \Gamma_{x,y^*,t}]'$ is deterministic. For any initial commitment $\{\boldsymbol{\Phi}_{-1}, \boldsymbol{\Gamma}_{-1}\}$, the central bank chooses $\{\boldsymbol{\Phi}_t, \boldsymbol{\Gamma}_t\}_{t=0}^{\infty}$. The initial commitment is self-consistent if

$$\boldsymbol{\Phi}_{-1} \overset{d}{=} \boldsymbol{\Phi}_t \sim \boldsymbol{\Phi}, \quad \forall t \geq 0, \tag{13}$$

$$\boldsymbol{\Gamma}_{-1} = \boldsymbol{\Gamma}_t = \boldsymbol{\Gamma}, \quad \forall t \geq 0, \tag{14}$$

which means that $\boldsymbol{\Phi}_{-1}$ and $\{\boldsymbol{\Phi}_t\}_{t=0}^{\infty}$ follow the same unconditional distribution, and $\boldsymbol{\Gamma}_{-1}$ and $\{\boldsymbol{\Gamma}_t\}_{t=0}^{\infty}$ are the same deterministic matrices. As the initial commitment is self-consistent, the policy with commitment becomes time-invariant.

To compute the equilibrium with the robustly optimal policy under the commitment device, we follow the solution method in Kwon and Miao (2019), which generalizes the Lagrangian game approach with conditionally linear policy commitment in Woodford (2010). The system of equations for the linearized equilibrium can be rewritten in the following form:

$$\begin{bmatrix} \boldsymbol{X}_{t+1} \\ \hat{\mathbb{E}}_t \boldsymbol{Y}_{t+1} \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{bmatrix} + \boldsymbol{B} i_t + \boldsymbol{C} \boldsymbol{\varepsilon}_{t+1} \tag{15}$$

where $\boldsymbol{X}_t \equiv [1, u_t, a_t, y_t^*]'$ is a vector of exogenous variables, $\boldsymbol{Y}_t \equiv [\pi_{H,t}, x_t]'$ is a vector of domestic inflation and the output gap which are non-predetermined variables, i_t is the nominal interest rate, which is the policy instrument, and $\boldsymbol{\varepsilon}_{t+1} \equiv [\varepsilon_{t+1}^u, \varepsilon_{t+1}^a, \varepsilon_{t+1}^y]'$ is a vector of shocks to the exogenous variables.

The minimax problem in (11) can be rewritten as the problem of choosing $\{\boldsymbol{\Phi}_t, \boldsymbol{\Gamma}_t\}_{t\geq 0}$, $\{m_t\}_{t\geq 1}$, and $\{i_t\}_{t\geq 0}$ for a given initial commitment $(\boldsymbol{\Phi}_{-1}, \boldsymbol{\Gamma}_{-1})$. The Lagrangian of the problem is given as follows:

$$\mathcal{L} = \mathbb{E}_{-1} \sum_{t=0}^{\infty} \beta^t \left\{ \frac{1}{2} \left(\pi_{H,t}^2 + \lambda_x \left(x_t - \bar{x} \right)^2 \right) - \theta m_{t+1} \ln m_{t+1} + \phi_t (\mathbb{E}_t m_{t+1} - 1) \right.$$

$$+[\boldsymbol{\mu}'_{X,t+1}, \boldsymbol{\mu}'_{Y,t}] \left(\begin{bmatrix} \boldsymbol{X}_{t+1} \\ \mathbb{E}_t m_{t+1} \boldsymbol{Y}_{t+1} \end{bmatrix} - \boldsymbol{A} \begin{bmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{bmatrix} - \boldsymbol{B} i_t \right) - \boldsymbol{C} \boldsymbol{\varepsilon}_{t+1} \Bigg\},$$

where $\beta^t \phi_t$ is the Lagrange multiplier for the constraint (1) and $\beta^t [\boldsymbol{\mu}_{X,t+1}, \boldsymbol{\mu}_{Y,t}]$ are vectors of the Lagrange multiplier for the system of equations (2), (3), (4), (8), (9), and (10).

Solving this problem consists of the following steps. First, the hypothetical malevolent nature chooses m_{t+1} to maximize the loss, which gives the solution for the worst-case belief m_{t+1}. Next, the policymaker chooses $\{\boldsymbol{X}_t, \boldsymbol{\Phi}_t, \boldsymbol{\Gamma}_t, i_t\}$ after substituting for the chosen solution of m_{t+1} in the Lagrangian. Since we are interested in self-consistent policy, we additionally impose $\boldsymbol{\Gamma}_t = \boldsymbol{\Gamma}$. To obtain the solution for $\boldsymbol{\Gamma}$, we start with a guess for $\boldsymbol{\Gamma}$ and solve the first-order conditions except for the first-order condition for i_t. The solution is then used to obtain a new value of $\boldsymbol{\Gamma}$ from the first-order condition for i_t. This process is repeated until we obtain convergence in the value of $\boldsymbol{\Gamma}$. We solve the resulting system of linear differential equations by using Klein (2000)'s method. The solution of the system takes the following state-space form of the law of motion of the state variables:

$$\begin{bmatrix} \boldsymbol{\varepsilon}_{t+1} \\ \boldsymbol{X}_{t+1} \\ \boldsymbol{\Phi}_t \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{C}_X \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\varepsilon}_{t+1}, \tag{16}$$

the policy rules for the non-predetermined variables:

$$\begin{bmatrix} i_t \\ \boldsymbol{\mu}_{Y,t} \\ \boldsymbol{\mu}_{X,t} \\ \mathbb{E}_t \boldsymbol{\mu}_{X,t+1} \end{bmatrix} = \boldsymbol{G} \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} \end{bmatrix}, \tag{17}$$

and the distorted belief measure, distorted forward-looking expectation, and the relative entropy in the worst-case scenario:

$$m_{t+1} = \exp\left(-\frac{1}{2} \theta^{-2} \boldsymbol{\mu}'_{Y,t} \boldsymbol{\Gamma} \boldsymbol{\Gamma}' \boldsymbol{\mu}_{Y,t} + \theta^{-1} \boldsymbol{\mu}'_{Y,t} \boldsymbol{\Gamma} \boldsymbol{\varepsilon}_{t+1} \right), \tag{18}$$

$$\mathbb{E}_t m_{t+1} \boldsymbol{Y}_{t+1} = \boldsymbol{\Phi}_t + \theta^{-1} \boldsymbol{\Gamma} \boldsymbol{\Gamma}' \boldsymbol{\mu}_{Y,t}, \tag{19}$$

$$\mathbb{E}_t m_{t+1} \ln m_{t+1} = \frac{1}{2} \theta^{-2} \boldsymbol{\mu}'_{Y,t} \boldsymbol{\Gamma} \boldsymbol{\Gamma}' \boldsymbol{\mu}_{Y,t}. \tag{20}$$

See Appendix A.1. for the entire procedure to solve the model.

3 Normative Analyses of the Robustly Optimal Monetary Policy

In this section, we investigate the normative features of the robustly optimal monetary policy. Specifically, we investigate the optimal policy coefficient of the domestic inflation and the output gap in equation (14), and impulse response functions of macroeconomic variables to exogenous shocks, subject to various degrees of concerns about distorted expectations and various structural parameters. Thereafter, we frequently use the acronym 'RE' when we refer to the model under $\theta^{-1} = 0$. First, we begin by claiming the following two propositions.

Proposition 3.1 *For any $\theta \in \mathbb{R}_+$, the robustly optimal policy calls for zero domestic inflation and output gap in response to domestic productivity shock a_t and world output shock y_t^*, i.e., $\Gamma_{\pi_H,a} = \Gamma_{x,a} = \Gamma_{\pi_H,y^*} = \Gamma_{x,y^*} = 0$.*

Proposition 3.2 *For any $\theta \in \mathbb{R}_+$, $\hat{\mathbb{E}}_t x_{t+1} = \mathbb{E}_t x_{t+1}$ in the equilibrium with the robustly optimal monetary policy.*

Proposition 3.1 implies that regardless of the central bank's concerns about distorted expectations, only the cost-push shock u_t generates a non-trivial reaction of $\pi_{H,t}$ and x_t in the equilibrium with the robustly optimal policy. The intuition is that the natural interest rate $\bar{rr}_t$ in (3) compromises the two shocks. The shock in the natural interest rate shifts the aggregate demand equation (2). It can be perfectly stabilized by the robustly optimal policy regardless of the potentially distorted expectation. The insight is that there are no trade-offs between stabilizing domestic inflation and the output gap. Thus, both can be simultaneously stabilized, which is also known as the divine coincidence. Proposition 3.1 does not imply that the natural rate shocks do not affect the economy. The natural rate shocks affect the natural level of domestic output $\tilde{y}_t$. Thus they affect the dynamics of the terms of trade and the nominal devaluation rate through equations (5) and (7). Figures C.14 - C.15 in Appendix C show the dynamic impulse response functions of the variables to the natural rate shocks.

Proposition 3.2 implies that in the equilibrium, the distorted expectations of the private agents that the central bank is concerned about is the expectation for domestic inflation, $\hat{\mathbb{E}}_t \pi_{H,t+1}$. The intuition is the following. The facts that (i) the optimal policy can be determined without the aggregate demand equation (2) and (ii) the aggregate supply equation (4) does not involve forward-looking expectations for the output gap motivate the central bank not to be concerned about distorted expectations for the output gap.

Based on the propositions, we focus in the next subsections on the robustly optimal policy with concerns about distorted expectations for domestic inflation in response to the cost-push shock.

3.1 Features of the Robustly Optimal Policy Coefficients

Figure 2 shows the robustly optimal policy coefficients under various degrees of concern about distorted expectations $\theta^{-1} \in \{0, 100, 500\}$, and its interaction with the standard deviations of the cost-push shock $\sigma_u \in [0, 0.05]$ and persistence of the shock $\rho_u \in [0, 1)$. For the other parameters, we set $\sigma = 1$, $\omega = 1$, $\varphi = 3$, $\beta = 0.99$, $\zeta = 0.75$, and $\alpha = 0.4$ following the benchmark calibration in Gali and Monacelli (2005).

The upper panels of the figure exhibit the optimal policy coefficients for domestic inflation, $\Gamma_{\pi_H,u}$. When $\theta^{-1} = 0$, the private expectations are RE. In this case, for any given level of persistence ρ_u, the policy coefficient $\Gamma_{\pi_H,u}$ is an increasing linear function of the standard deviation of the cost-push shock, σ_u, which implies that the *certainty equivalence* is applied in the optimal monetary policy.

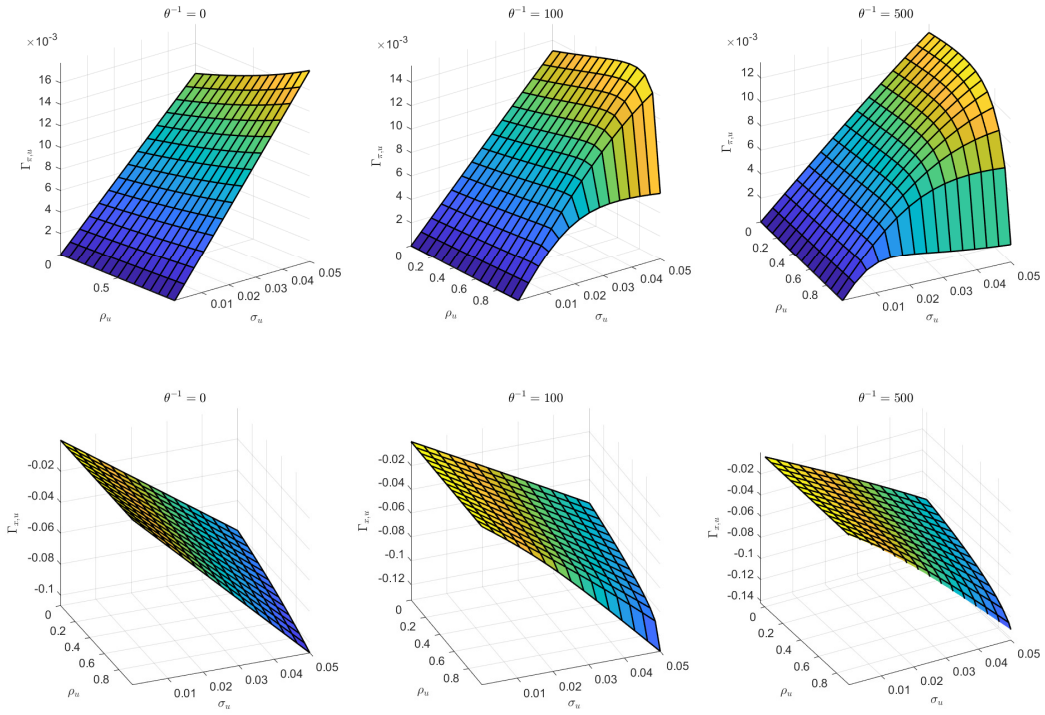
If $\theta^{-1} > 0$, however, then the central bank is concerned about the distorted expectations of private agents. The central bank's concern causes a breakdown of the certainty equivalence principle in the optimal monetary policy, as discussed in Woodford (2010). For example, when $\theta^{-1} = 100$, $\Gamma_{\pi_H,u}$ is an increasing but concave function of σ_u, which implies that the policy becomes more *conservative* than the policy under RE. This implies that as σ_u increases, the central bank becomes more reluctant in responding to $\Gamma_{\pi_H,u}$ because of its concerns about distorted expectations. Furthermore, as the persistence of the cost-push shock ρ_u increases, the decrease in the rate of the increase in $\Gamma_{\pi_H,u}$ to σ_u increases, which implies that the central bank becomes increasingly conservative in responding $\Gamma_{\pi_H,u}$ when the persistence of the shock increases. Remarkably, when $\theta^{-1} = 500$, $\Gamma_{\pi_H,u}$ begins to even decrease above a certain point σ_u when ρ_u is sufficiently high. The central bank now decides to decrease its response to $\Gamma_{\pi_H,u}$ given its serious concerns about distorted expectations.

The lower panels in Figure 2 exhibit the optimal policy coefficient for the response of the output gap $\Gamma_{x,u}$ in the models under θ^{-1} and its interaction with $\{\sigma_u, \rho_u\}$. The aggregate supply equation (4) implies that given domestic inflation $x_{H,t}$, the worst-case distorted expectation $\hat{\mathbb{E}}_t \pi_{H,t+1}$, and the shock u_t, the equilibrium output gap x_t is determined as follows:

$$x_t = \frac{\pi_{H,t} - \beta \hat{\mathbb{E}}_t \pi_{H,t+1} - u_t}{\kappa}, \tag{21}$$

which indicates that $\Gamma_{x,u}$ depends on the gap between the current domestic inflation and the distorted expectation for future inflation $\pi_{H,t} - \beta \hat{\mathbb{E}}_t \pi_{H,t+1}$, the cost-push shock u_t, and

Figure 2: Robustly Optimal Policy Coefficients Conditional on Various $\theta^{-1}, \sigma_u, \rho_u$

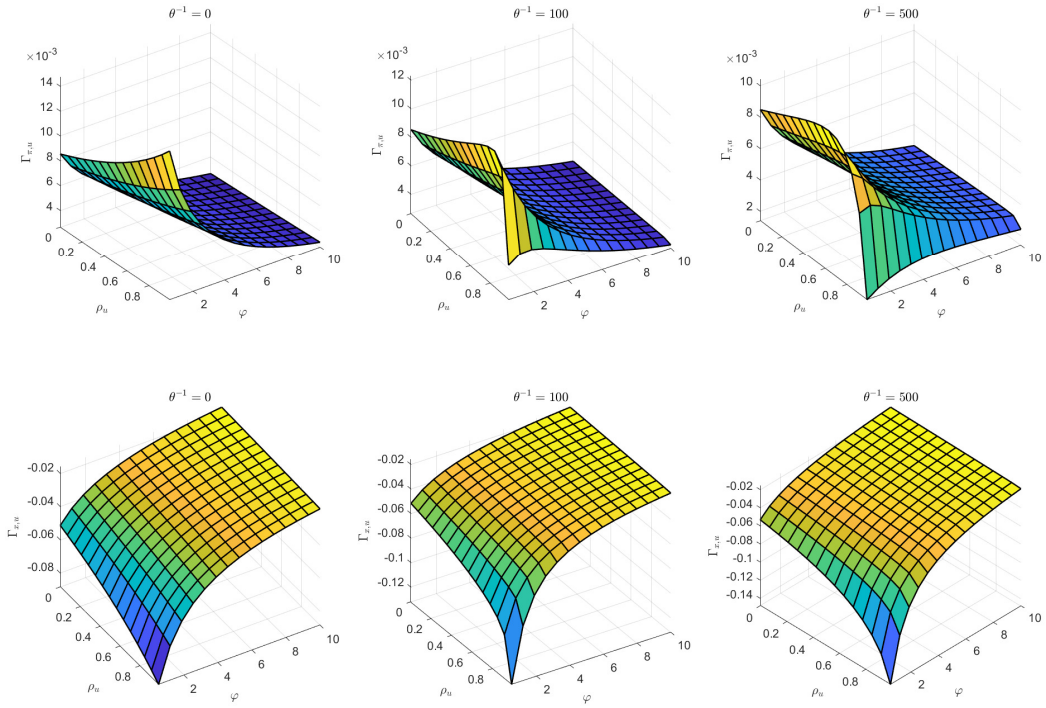


the slope of the aggregate supply curve κ. As in the case of $\Gamma_{\pi_H,u}$, when $\theta^{-1} = 0$, the private expectations are RE. At any level of persistence ρ_u, $\Gamma_{x,u}$ is a linear function of σ_u since certainty equivalence is applied. In addition, the policy $\Gamma_{x,u}$ is a decreasing function of σ_u, since an increase in σ_u increases the size of u_t when the shock occurs, and it negatively affects x_t by the aggregate supply relation (21).

When $\theta^{-1} > 0$, in contrast, the concerns about distorted expectations arise. The coefficient $\Gamma_{x,u}$ is not a linear function of σ_u, and as ρ_u increases, $\Gamma_{x,u}$ decreases further than it decreases in the case of certainty equivalence. The mechanism is as follows. As we see in the upper panels, as σ_u and ρ_u increase, the motivation for the conservative response of $\Gamma_{\pi_H,u}$ becomes stronger. The motivation is to give a signal to private agents that the central bank does not have plans to significantly change domestic inflation to maintain the gap $\pi_{H,t} - \beta\hat{\mathbb{E}}_t\pi_{H,t+1}$ to a small and stable level. As ρ_u and σ_u increase, the central bank wants to get $\pi_{H,t} - \beta\hat{\mathbb{E}}_t\pi_{H,t+1}$ increasingly smaller. With equation (21), these result in a larger decrease in $\Gamma_{x,u}$ than in the case of RE. When $\theta^{-1} = 500$, the pattern becomes quantitatively stronger.

The upper panels in Figure 3 show the $\Gamma_{\pi_H,u}$ under θ^{-1} and its interaction with $\{\varphi, \rho_u\}$ (we fix $\sigma_u = 0.02$). When φ increases, the slope of the aggregate supply curve becomes

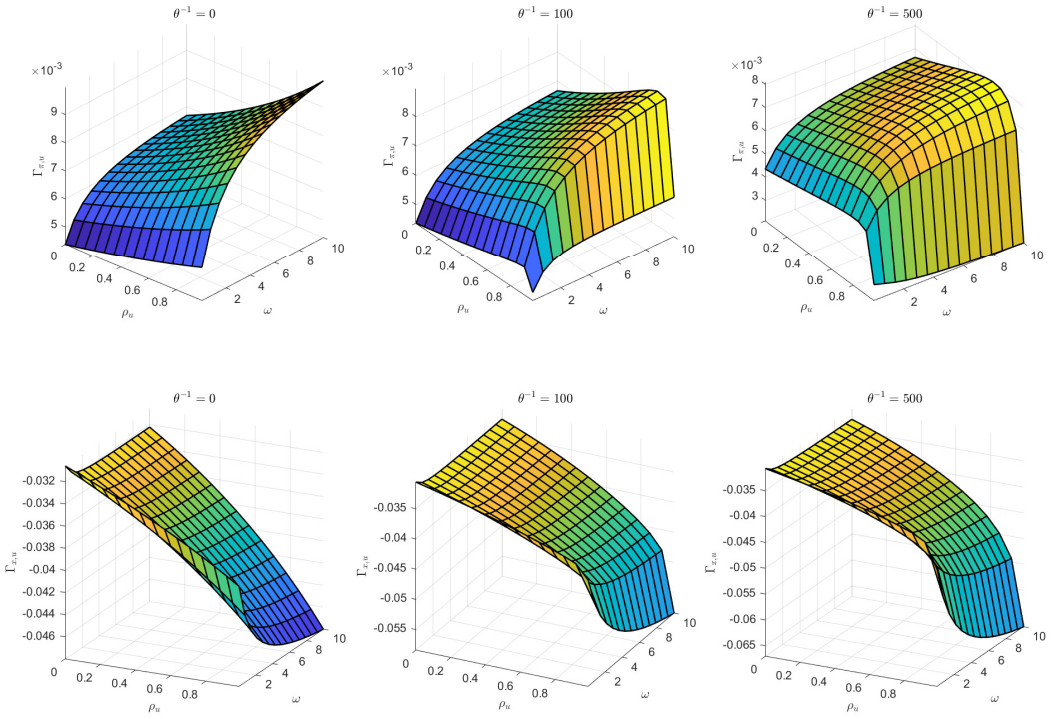
Figure 3: Robustly Optimal Policy Coefficients Conditional on Various $\theta^{-1}, \varphi, \rho_u$



steeper, which implies that the propagation effect of monetary policy becomes weaker. Thus at all degrees of concerns about distorted expectations θ^{-1} and all levels of persistence of the cost-push shock ρ_u, the policy coefficient $\Gamma_{\pi_H,u}$ exhibits a low response when φ is high. When φ is small, i.e., the aggregate supply curve is flatter, the policy coefficient becomes substantially different along with the concerns about distorted expectations. If $\theta^{-1} = 0$, then $\Gamma_{\pi_H,u}$ increases when φ decreases. Moreover, $\Gamma_{\pi_H,u}$ increases when ρ_u increases. This implies that when the private expectations are RE, the optimal monetary policy calls for aggressive monetary policy (higher $\Gamma_{\pi_H,u}$) when the aggregate supply curve is flatter and when the cost-push shock is more persistent. If $\theta^{-1} = 100$, however, then $\Gamma_{\pi_H,u}$ begins to decrease when φ becomes sufficiently small and ρ_u becomes sufficiently high. The pattern of the decrease in $\Gamma_{\pi_H,u}$ becomes quantitatively stronger when $\theta^{-1} = 500$. The results have the following implications. Suppose that the cost-push shock is substantially persistent and that the aggregate supply curve is substantially flat. In this case, the central bank decides to restrain its actions and conducts conservative monetary policy concerning distorted expectations.

The lower panels in Figure 3 show $\Gamma_{x,u}$ under the same variations of the parameters as in the upper panels. As we observed in the upper panels, a higher φ yields a higher κ, which increases the denominator of the right-hand side of equation . Thus, $\Gamma_{x,u}$ uniformly increases

Figure 4: Robustly Optimal Policy Coefficients Conditional on Various $\theta^{-1}, \omega, \rho_u$



toward 0 when φ increases for any given ρ_u and θ^{-1}. When $\theta^{-1} = 0$ (RE), $\Gamma_{x,u}$ decreases as ρ_u increases for any given φ. The reason is because the gap $\pi_{H,t} - \beta \mathbb{E}_t \pi_{H,t+1}$ decreases when the persistence of the shock increases, since the expected path of domestic inflation also has a higher serial correlation. When $\theta^{-1} = 100$, the gap $\pi_{H,t} - \beta \hat{\mathbb{E}}_t \pi_{H,t+1}$ becomes even smaller because of the conservative response of the $\Gamma_{\pi_H,u}$ caused by the concerns about distorted expectations. Thus $\Gamma_{x,u}$ decreases more than the case of RE, when ρ_u increases. The pattern becomes quantitatively stronger when $\theta^{-1} = 500$.

The upper panels in Figure 4 show $\Gamma_{\pi_H,u}$ under θ^{-1} and its interaction with $\{\omega, \rho_u\}$ (we fix $\sigma_u = 0.02$ and $\varphi = 1$). The parameter ω governs the slope of the aggregate demand curve $-\frac{1}{\sigma_\alpha}$, since $\sigma_\alpha \equiv \frac{\sigma}{1+\alpha(\omega-1)}$. A higher ω implies a steeper slope of the aggregate demand curve. The parameter ω also affects the aggregate supply curve and the slope κ_α becomes flatter as ω increases. Thus, an increase in ω strengthens the propagation effect of the monetary policy. If $\theta^{-1} = 0$ (RE), then $\Gamma_{\pi_H,u}$ increases when ω increases. If $\theta^{-1} = 100$, then the increase in $\Gamma_{\pi_H,u}$ becomes sluggish when ρ_u is sufficiently high. When $\theta^{-1} = 500$, $\Gamma_{\pi_H,u}$ actually decreases as ω increases if ρ_u is substantially persistent. The essential insight is similar to the case in Figure 3. A steeper aggregate demand curve and a flatter aggregate supply curve caused by a higher ω strengthen the effect of monetary policy. As there are concerns about

distorted expectations, the central bank decides to conduct conservative monetary policy by adopting a smaller $\Gamma_{\pi_H,u}$ than the one under RE.

The lower panels in Figure 4 show $\Gamma_{x,u}$ under the same variations of the parameters as in the upper panels. As the higher ω causes a flatter κ, the $\Gamma_{x,u}$ uniformly decreases when ω increases, for any given ρ_u and θ^{-1}. When $\theta^{-1} = 0$ (RE), at any fixed ω, the $\Gamma_{x,u}$ decreases as ρ_u increases. Similar to the previous case, the reason is because the gap $\pi_{H,t} - \beta\mathbb{E}_t\pi_{H,t+1}$ decreases with the higher persistence of the shock. When $\theta^{-1} = 100$, the declining patterns become stronger because of the concerns about distorted expectations. When ρ_u becomes higher, the decrease in $\Gamma_{x,u}$ becomes more drastic than in the case of RE. The reason is because the gap $\pi_{H,t} - \beta\hat{\mathbb{E}}_t\pi_{H,t+1}$ decreases as a consequence of the central bank's conservative monetary policy. The case of $\theta^{-1} = 500$ strengthens the pattern.

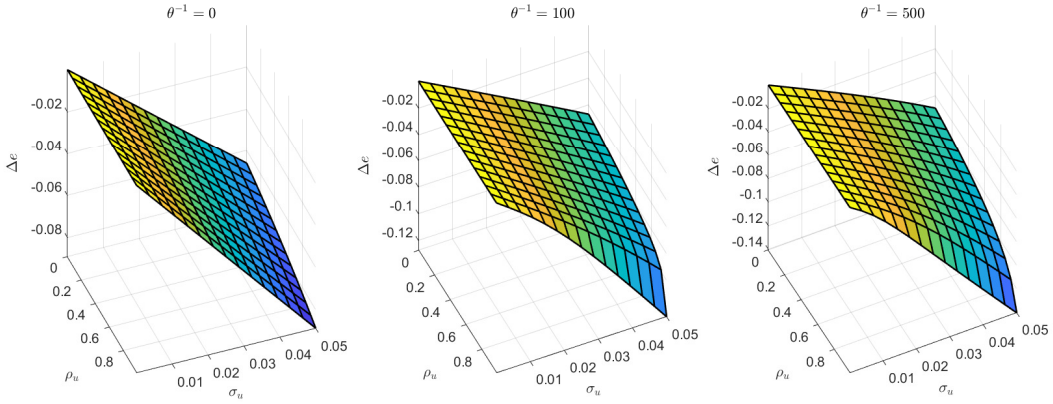
3.2 Features of the Initial Response of the Nominal Devaluation Rate

How does the robustly optimal monetary policy affect the nominal exchange rate? Using the model solutions of $\{\pi_{H,t}, x_t, \tilde{y}_t\}$ and exogenous process y_t^*, we obtain the solution for the devaluation rate Δe_t from (5) and (7) in the equilibrium.

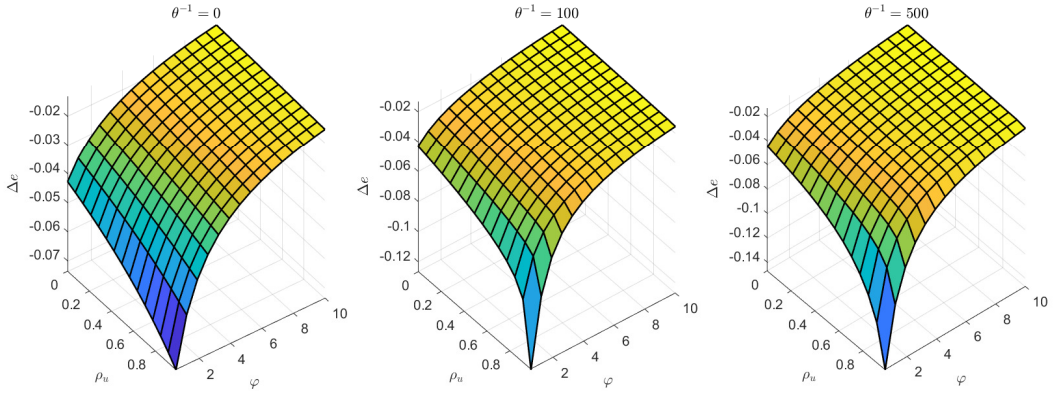
Figure 5 shows the initial responses of the nominal devaluation rate Δe_t at the time when the cost-push shock occurs, in the models with $\theta^{-1} \in \{0, 100, 500\}$ and their interactions with other various structural parameters. Here, we fix the other parameters for the panels in (a), (b), and (c) to be the same as those in the cases of Figures 2, 3, and 4, respectively. The panels in (a) exhibit the initial responses under various θ^{-1} and its interactions with (σ_u, ρ_u). When $\theta^{-1} = 0$ (RE), Δe_t decreases (initial appreciation) as σ_u increases, at any level of persistence ρ_u. When $\theta^{-1} = 100$, the initial appreciation subject to the increase in σ_u accelerates more than the initial appreciation under RE. When $\theta^{-1} = 500$, the pattern of the acceleration strengthens. Thus, the model shows a far larger initial appreciation than the former two models. Moreover, the gaps of the initial response of Δe_t between the models with different degrees of distorted expectations become more dramatic when the shock is more persistent. All three models with different θ^{-1} values show a larger initial appreciation subject to σ_u when ρ_u is higher. However, the initial responses of Δe_t in the model under RE are quite proportional to σ_u, whereas when $\theta^{-1} > 0$, they become exponential as σ_u increases. The qualitative features of the initial responses are close to the features of the robustly optimal policy coefficients of the output gap $\Gamma_{x,u}$ in Figure 2. Equations (5) and (7) imply that the response of Δe_t to the cost-push shock is affected by the responses of $\pi_{H,t}$ and $\sigma_\alpha \Delta x_t$, and the effect of Δx_t quantitatively dominates the effect of $\pi_{H,t}$. As we can see in Figure 2, the reasons are because $|\Gamma_{x,u}|$ are quantitatively larger than $|\Gamma_{\pi_H,u}|$, and $\Gamma_{\pi_H,u}$

Figure 5: Initial Responses of the Devaluation Rate to the Cost-Push Shock

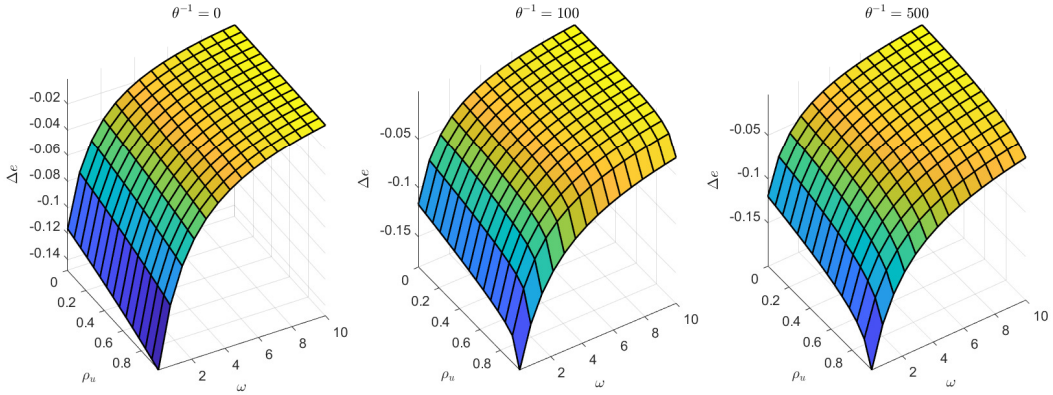
(a) Conditional on Various $\theta^{-1}, \sigma_u, \rho_u$



(b) Conditional on Various $\theta^{-1}, \varphi, \rho_u$



(c) Conditional on Various $\theta^{-1}, \omega, \rho_u$



behaves sluggishly to an increase of σ_u when $\theta^{-1} > 0$.

The panels in (b) show the responses under θ^{-1} and their interactions with (φ, σ_u). We observe that as θ^{-1} increases, φ is smaller, and ρ_u is more persistent, there are more initial appreciations. The initial responses of Δe_t are similar to the policy coefficient $\Gamma_{x,u}$ in Figure 3. The insight is similar to the case in the panels in (a). The panels in (c) show the responses under θ^{-1} and its interaction with (ω, σ_u). Similar to the panels in (b), there are more initial appreciations as θ^{-1} increases, ω is smaller, and ρ_u is more persistent. However, the responses are quite different from the policy coefficient $\Gamma_{x,u}$ in Figure 4, in which $\Gamma_{x,u}$ does not necessarily decrease when ω decreases. The reason is because $\sigma_\alpha \equiv \frac{\sigma}{1+\alpha(\omega-1)}$ increases as ω decreases. Thus, the impact of $\Gamma_{x,u}$ on Δe_t is amplified as ω decreases.

The results in Figure 5 indicate that the central bank's conservative monetary policy as a response to the concerns about distorted expectations causes a more aggressive initial response of the nominal exchange rate. The general intuition can be summarized as follows. The cost-push shock creates trade-offs between domestic inflation and the output gap in all models. When $\theta^{-1} > 0$, the initial response of the output gap becomes larger than that under RE because of the sluggish response of domestic inflation. Since the nominal devaluation is determined by equation (7), the movements of domestic inflation and the change of the terms of trade both matter. Equation (5) indicates that the terms of trade is a linear function of the output gap. The magnitude of the response of the output gap outweighs the response of domestic inflation for a broad range of parameters, and the magnitude increases as θ^{-1} increases. Consequently, the nominal devaluation shows a larger initial response than that under RE when $\theta^{-1} > 0$.

3.3 Robustly Optimal Dynamics

In this subsection, we investigate the dynamic features of the equilibrium with the robustly optimal monetary policy. We begin by claiming the following proposition.

Proposition 3.3 *In the equilibrium with the robustly optimal monetary policy, the dynamics of domestic inflation is determined by the following equation:*

$$\pi_{H,t} = \rho_{\pi_H}\pi_{H,t-1} - \Gamma_{\pi_H,u}\varepsilon_t^u - \rho_\pi(1-\rho_u)\sigma_u \sum_{j=0}^{\infty} \rho_u^j \varepsilon_{t-j-1}^u, , \tag{22}$$

where $\rho_{\pi_H} \in (0,1)$ *increases as* θ^{-1} *increases.*

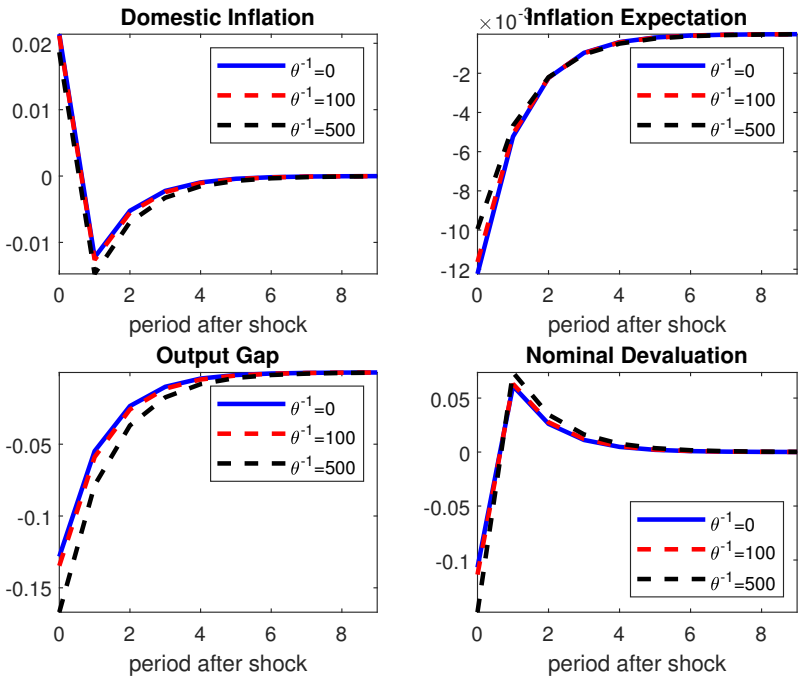
Proposition 3.3 means the following. As the central bank has more concerns about distorted expectations, the coefficient ρ_{π_H} increases toward one, so the current domestic inflation

becomes more affected by past domestic inflation. In addition, if the cost-push shock has nonzero persistence, then the current domestic inflation is more affected by the weighted average of the history of past cost-push shocks as the central bank has more concerns about distorted expectations. These actions imply that the rise of the central bank's concerns results in more history-dependent movements in domestic inflation than those in the rational expectation benchmark. The result recalls the one reached in Woodford (2010) in the closed-economy framework, and proposition 3.3 contains a more general result because of the consideration of the shock persistency.

Figure 6 shows the dynamic impulse response functions of domestic inflation, the (worst-case) private expectations for domestic inflation, the output gap, and the nominal devaluation rate to one standard deviation cost-push shock. In the exercise, we set $\rho_u = 0$ (no persistence), $\sigma_u = 0.05, \varphi = 0.5$ and $\omega = 1$. The impulse response functions of the variables exhibit optimal expected paths that are induced by the central bank's intervention after the occurrence of the shock. When $\theta^{-1} = 0$ (blue solid line), the domestic inflation initially jumps in response to the shock but then shifts to a negative number (deflation) in the following period and then gradually converges to zero over time. The private expectations are fully rational and are model-consistent. Thus, the dynamics of next-period domestic inflation and inflation expectations are consistent. The output gap initially shifts to a negative number and then gradually converges to zero. The impulse response functions show optimal inflation-output gap trade-offs to stabilize private agents' rational expectations. The nominal devaluation rate initially shifts to a negative number (initial appreciation), jumps to a positive number (devaluation) in the following period, and then gradually converges to zero. The initial response of the devaluation rate is mainly caused by the initial response of the output gap. The dynamic responses of the devaluation rate, in contrast, are mainly related to the lagged differences of the responses of the output gap, as implied in (5) and (7).

When $\theta^{-1} = 100$ (red dashed line), domestic inflation jumps less than that under RE because of the smaller $\Gamma\pi_H, u$. It also maintains a lower level than under RE over the dynamics because of the higher ρ_{π_H} in equation (22). This implies that the economy has a lower domestic price level than that under RE. The output gap initially drops more than under RE and then converges to zero over time. These actions show robustly optimal dynamic trade-offs that are implemented by the central bank, which seeks to stabilize private agents' worst-case expectations. When $\theta^{-1} > 0$, the private expectations become worst-case expectations, which are no longer model-consistent. The dynamic responses of the worst-case expectation maintain somewhat milder behavior than those under RE in terms of magnitude. This implies that the central bank puts more effort into stabilizing private expectations to manage the worst-case outcome caused by distorted expectations. These results are in line

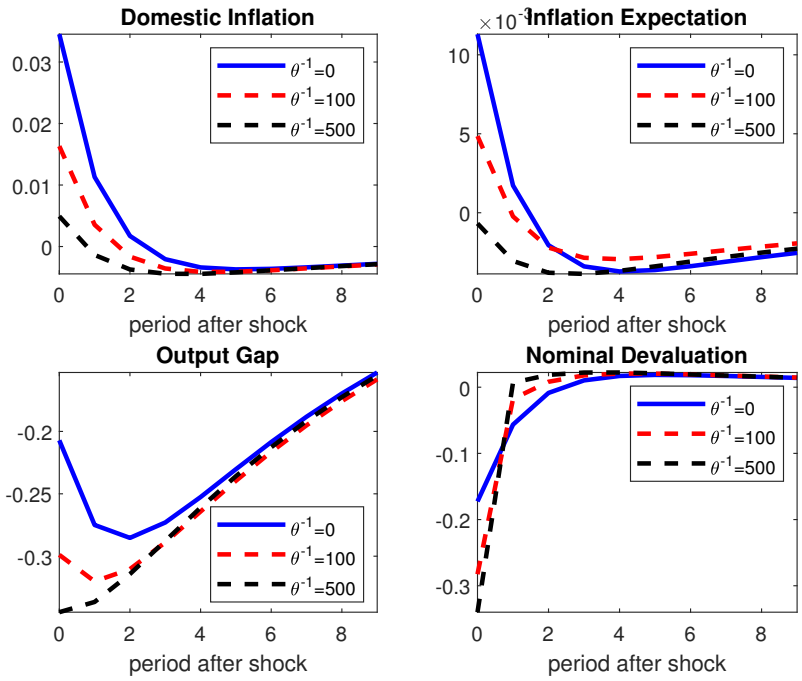
Figure 6: Dynamic Responses of Macroeconomic Variables to the Cost-Push Shock: $\rho_u = 0$



with the prediction in the closed-economy environment in Woodford (2010). The nominal devaluation rate, in contrast, behaves in a different way. The nominal devaluation rate initially drops below that under RE and jumps in the following period more than that under RE. Then, the nominal devaluation rates maintain higher levels than those under RE over the dynamics. The dynamics of the nominal devaluation rates are also closely linked to the initial response and the lagged differences of the dynamic responses of the output gap. The patterns of the dynamic responses of the variables become quantitatively more substantial when $\theta^{-1} = 500$ (black dashed line).

Figure 7 shows the dynamic impulse response functions of the same variables when the cost-push shock is highly persistent, $\rho_u = 0.9$. When $\theta^{-1} = 0$ (RE), the optimal monetary policy calls for a gradual and smooth decrease in the responses of domestic inflation over time after the initial jump. They are qualitatively different from the responses when the persistence of the shock is mild, as shown in Figure 6. These also imply different optimal dynamic trade-offs. The output gap now shows hump-shaped dynamic responses. After the initial drop, the output gap drops further for the two following periods. The reason is because the output gap is determined by equation (21): (i) the gap between domestic inflation and its expectation is smaller than in the case in Figure 6, and (ii) the cost-push shock remains high because of its high persistence. The nominal devaluation rate initially

Figure 7: Dynamic Responses of Macroeconomic Variables to the Cost-Push Shock: $\rho_u = 0.9$



drops and then gradually converges to zero. Unlike the case in Figure 6, the devaluation rate does not jump to a positive number after the initial drop. The reason is because the lagged difference Δx_t remains negative for the following two periods by the hump-shaped impulse response functions.

When $\theta^{-1} = 100$, the initial responses of the domestic inflation response are substantially less than those under RE because of the substantially smaller $\Gamma\pi_H, u$. In addition, it maintains a significantly lower level than under RE over the dynamics because of higher ρ_{π_H} and ρ_u in equation (22). The worst-case inflation expectation also shows milder dynamic responses than those under RE. These mild responses (caused by concerns about distorted expectations) are significantly prominent over the ones in Figure 6. As discussed in Figures 2-4, the essential mechanism of the results is that the higher persistence of the shock strengthens the central bank's motivation to conduct more conservative monetary policy when it is concerned about distorted expectations. The output gap has a larger initial drop than RE, drops slightly further in the following period, and then converges to zero. The output gap shows weaker hump-shaped responses than RE. Consequently, the devaluation rate initially drops more than RE, jumps to a negative number close to zero in the following period, and then converges to zero over time. When $\theta^{-1} = 500$, the milder responses of domestic inflation and its worst-case expectation become more prominent. The output gap

has a larger initial drop and does not have hump-shaped responses in the following periods. Consequently, the devaluation rate has a larger initial drop, jumps to a positive number near zero, and then shows a mild inverse-U shape over time until it converges to zero. As concerns about distorted expectations rise, the nominal exchange rate shows milder dynamics over time after an aggressive initial response.

4 Bayesian Inference

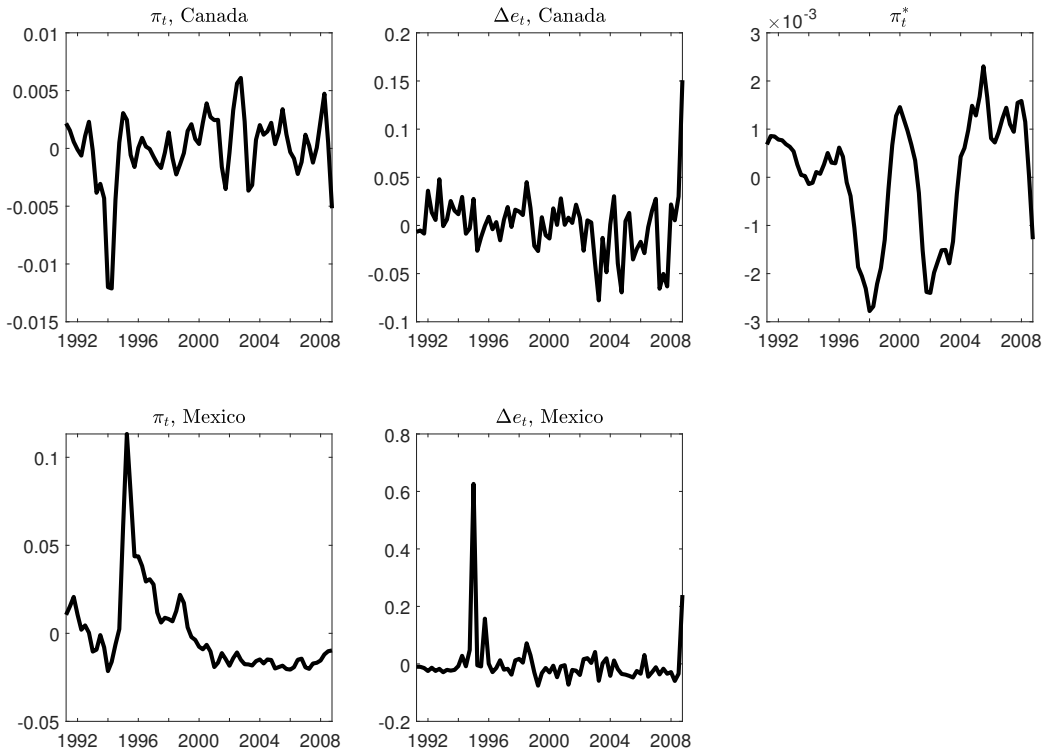
In this section, we perform empirical analyses using the models of the robustly optimal policy. The crucial questions regarding the models are (i) what the actual data tell about the parameters of the model, including the central bank's concerns about distorted expectations, and (ii) how the estimated models of optimal monetary policy predict actual monetary policies. We consider two classes of models to answer these questions. One is the RE model with $\theta^{-1} = 0$, and the other is the model with distorted expectations, in which θ^{-1} is an estimable parameter. Throughout this section, we use the acronym 'NRE' when we refer to the model in which θ^{-1} is an estimable parameter. Estimating the models of the robustly optimal monetary policy means that we interpret the actual data as an outcome interaction with the optimal monetary policy from the policymakers. All models are estimated by using time-series data and the likelihood-based Bayesian method.

4.1 Data

We consider Canada and Mexico for our empirical analyses. On the one hand, both are considered to be typical small open economies. They are located in North America, have the United States as their main trade partner, and are OECD member countries (Canada: joined in 1961, Mexico: joined in 1994). On the other hand, Canada is considered to be a typical advanced country, whereas Mexico is regarded as a typical emerging market counterpart. The business cycle fluctuations of Mexico have been substantially more volatile than those of Canada. Furthermore, the cyclical patterns of prices and quantities, such as inflation rates, exchange rates, and cross-border capital flows, have been significantly different between the two countries. What the data tell us about the models of the robustly optimal policy for the countries is a question of interest.

We use quarterly data on the CPI inflation rate, nominal devaluation rate, and US CPI inflation rate as observables. To construct the observables, we use quarterly CPI price indices and nominal exchange rates (vis-á-vis the US dollar). All price indices are seasonally adjusted by using the program package X13-ARIMA-SEATS. The time spans of the CPI inflation rate and the nominal devaluation rate are Q2:1991–Q4:2008 for Canada and Q1:1991–Q4:2008

Figure 8: Time Series Observations



for Mexico. The US CPI inflation rate is used for the overlapping periods for each country. Figure 8 exhibits the time-series observations for CPI inflation rates π_t, nominal devaluation rates Δe_t, and US CPI inflation π_t^* over the sample periods. All series are demeaned. It is observed that the CPI inflation and nominal devaluation rates do not show significant comovements in Canada. In Mexico, however, there are strong comovements between the two series during the Mexican Peso Crisis (1994-1996). The CPI inflation rate in Mexico was significantly affected by the exchange rate pass-through during the episode of large devaluation.

4.2 The Linear State Space System

We reproduce the system that governs equilibrium solutions of the model. We denote the vector of state variables as $\boldsymbol{\xi}_t \equiv [\boldsymbol{\varepsilon}_t', \boldsymbol{X}_t', \boldsymbol{\Phi}_{t-1}']'$ and denote $\boldsymbol{S}^{\varepsilon} \equiv [\boldsymbol{I}, \boldsymbol{C}_X', \boldsymbol{0}]'$, where $\boldsymbol{I}$ and $\boldsymbol{0}$ are identity and zero matrices, respectively. Equation (16) can be rewritten as

$$\boldsymbol{\xi}_{t+1} = \boldsymbol{H}\boldsymbol{\xi}_t + \boldsymbol{S}^{\varepsilon}\boldsymbol{\varepsilon}_{t+1}.$$

Similarly, we denote the set of non-predetermined variables as $\boldsymbol{\chi}_t \equiv [i_t, \boldsymbol{\mu}'_{Y,t}, \boldsymbol{\mu}'_{X,t}, \mathbb{E}_t\boldsymbol{\mu}'_{X,t+1}]'$, so that equation (17) can be rewritten as

$$\boldsymbol{\chi}_t = \boldsymbol{G}\boldsymbol{\xi}_t.$$

The time-series observables are used to identify some parameters of interest in the system of equations that represent equilibrium solutions of the model. The equations for aggregate demand and aggregate supply, (2) and (4), fully describe the equilibrium determination of domestic inflation and the output gap, subject to distorted beliefs and stochastic shocks. There are some practical limitations in measuring the pure domestic inflation rate $\pi_{H,t}$ and output gap x_t for the countries. However, we can indirectly map the data to the two model variables by using observables on domestic CPI inflation rates, nominal devaluation rates, and US CPI inflation rates. The CPI inflation π_t of a domestic country in the model is determined as

$$\pi_t = \pi_{H,t} + \alpha\Delta s_t, , \tag{23}$$

and the nominal devaluation rate Δe_t is determined by equation (7), where we now use series for the US CPI inflation rate as proxies for inflation in the rest of the world π^*_t. The process of $\tilde{y}^*_t$ is exogenously given in the model with calibrated parameters ρ_{y^*} and σ_{y^*}. Thus, we use observations on π_t and $\Delta\tilde{e}_t$ such that

$$\Delta\tilde{e}_t \equiv \Delta e_t + \pi^*_t, \tag{24}$$

with equations (7) and (23) yields indirect observations for $\pi_{H,t}$ and Δx_t for the model counterparts. The variable $\Delta\tilde{e}_t$ represents the inflation rate of the rest of the world measured in domestic currency. Note that the magnitude of fluctuations of US CPI inflation π^*_t in Figure 8 is much milder than the ones of Δe_t. This implies that the fluctuations of $\Delta\tilde{e}_t$ of the two countries mostly come from the fluctuations of Δe_t.

Since the set of variables needed to express π_t and $\Delta\tilde{e}$ as solutions to the equilibrium also features their lagged counterparts, the vector $[\boldsymbol{\xi}'_t, \boldsymbol{\xi}'_{t-1}]'$ should be used as the state variables in the system of equations. The measurement equations for the observables with the state variables (accounting for measurement errors) are given by:

$$\begin{bmatrix} \pi_t \\ \Delta\tilde{e}_t \end{bmatrix} = \tilde{\boldsymbol{G}} \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_{t-1} \end{bmatrix} + \begin{bmatrix} \sigma^{me}_{\pi} & 0 \\ 0 & \sigma^{me}_{\Delta\tilde{e}} \end{bmatrix} \cdot \begin{bmatrix} \epsilon^{me,\pi}_t \\ \epsilon^{me,\Delta\tilde{e}}_t \end{bmatrix}, \tag{25}$$

Table 1: Pre-Set Parameters

Parameter	Canada	Mexico
σ	1	
ζ	0.75	
ρ_{y^*}	0.92	
$\sigma_{y^*} \times 100$	0.55	
β	0.99	0.975
α	0.4	0.248

where $[\epsilon_t^{me,\pi}, \epsilon_t^{me,\Delta\tilde{e}}]'$ are orthogonal standard normal random variables and $[\sigma_\pi^{me}, \sigma_{\Delta\tilde{e}}^{me}]'$ are the standard deviations of the measurement errors on the observables for π_t and $\Delta\tilde{e}$, respectively.

The transition equations for the evolution of the state variables are given by

$$\begin{bmatrix} \boldsymbol{\xi}_{t+1} \\ \boldsymbol{\xi}_t \end{bmatrix} = \tilde{\boldsymbol{H}} \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_{t-1} \end{bmatrix} + \boldsymbol{\nu}_{t+1}. \tag{26}$$

Thus, equations (25) and (26) complete the description of the linear dynamic state space system. The details for $\boldsymbol{H}, \boldsymbol{S}^\varepsilon, \boldsymbol{G}, \tilde{\boldsymbol{H}}, \tilde{\boldsymbol{G}}$, and $\boldsymbol{\nu}_{t+1}$ are presented in Appendix B.

4.3 Pre-Set Parameters

Some of the parameters of the models are preset before the Bayesian estimation. These parameter values directly follow the previous literature and use the moments in the data without referring to the Bayesian estimation. These parameters can further be classified into the two subgroups: the structural parameters that have an economic interpretation $[\sigma, \zeta, \beta, \alpha]'$ and the parameters that govern the transition process of the world output shock $[\rho_{y^*}, \sigma_{y^*}]'$. The values of these parameters are listed in Table 1.

The parameter σ captures the curvature of the utility function of the household and is assumed to be 1, which results in a log utility function for all the countries, as in Gali and Monacelli (2005). The Calvo-Yun parameter of price stickiness, ζ, is assumed to be 0.75 for all countries, which signifies an average period of 1 year between price adjustments. For the rest of the world's output, we use the output series of the US and estimate an AR(1) regression. This yields a persistence, ρ_{y^*}, of 0.9261 and a standard deviation, σ_{y^*}, of 0.0055. The discount factor, β, is calibrated separately for each country. The value of β is chosen to match the average interest rate in the corresponding country over the sample period. Mendoza (1991) estimates the annual country interest rate for developed countries as 4% on average. Uribe and Schmitt-Grohé (2017) estimate the annual country interest rates of

Table 2: Prior Distributions of Parameters for Estimation

Parameter	Support	Distribution	Para (1)	Para (2)	[mean, std]
θ^{-1}	$[0, 10^{10}]$	Uniform	0	10^{10}	$[5 \cdot 10^9, 2.88 \cdot 10^4]$
ω	$\mathbb{R}^+$	Gamma	2.0	2.0	[4.0, 2.82]
φ	$\mathbb{R}^+$	Gamma	2.0	2.0	[4.0, 2.82]
ρ_u	$[0, 1)$	Beta	2.0	3.0	[0.4, 0.2]
ρ_a	$[0, 1)$	Beta	2.0	3.0	[0.4, 0.2]
$\sigma_u \times 100$	$\mathbb{R}^+$	Gamma	0.25	2.0	[0.5, 1.0]
$\sigma_a \times 100$	$\mathbb{R}^+$	Gamma	0.25	2.0	[0.5, 1.0]

emerging market countries as 10% on average. Thus, we set the value of β for Canada as 0.99, whereas we set it to 0.975 for Mexico. The home-bias parameter, α, is also calibrated separately for each country. It is calibrated to match the average import-GDP ratio over the time series. Thus, α is 0.4 for Canada and 0.248 for Mexico, which signifies that imports represent 40% of GDP for Canada and 25% of GDP for Mexico.

4.4 Estimated Parameters

We estimate the remaining seven structural parameters $\mathbf{\Theta} \equiv [\theta^{-1}, \omega, \varphi, \rho_u, \sigma_u, \rho_a, \sigma_a]'$ and two nonstructural parameters $[\sigma_\pi^{me}, \sigma_{\Delta e}^{me}]'$ by using the likelihood-based Bayesian method.

Table 2 provides the details of the prior distribution of the seven structural parameters. To account for the possibility of the RE, the minimum value of θ^{-1} is 0. The maximum value of θ^{-1} is 10^{10}, which implies a sizable departure from the RE. The prior distribution of θ^{-1} is assumed to be uniform between these lower and upper bounds. For the parameter ω, we allow for a broad range while Gali and Monacelli (2005) focuses on a special case of $\omega = 1$. The prior distribution of ω follows a Gamma (2,2) distribution, resulting in a mean of 4 and a standard deviation of 2.82. The Frisch elasticity of labor supply in the model is given as φ^{-1}. Given that many microeconomic estimates of the Frisch elasticity lie between 0.3 and 0.5, while many macroeconomists use an estimate between 2 and 4, we allow for a wide range of values for φ, assuming that φ follows a Gamma (2,2) distribution. The persistence of shock processes is assumed to follow a Beta (2,3) distribution, with a support of $[0, 1]$, which allows for a mean of 0.4 and a standard deviation of 0.2. The one hundred times of the standard deviations of the shock processes are assumed to follow a Gamma (0.25,2) distribution, with a support of $(0, \infty)$. This allows for a mean and standard deviation of 0.5 and 1.0, respectively. Finally, for the standard deviation parameters of the measurement errors of the observables, we impose uniform distributions with the maximum supports to be 25% of the standard deviation of each observable.

We obtain draws from the posterior distribution of the estimated parameters $\mathbf{\Theta}$ condi-

Table 3: Posterior Distribution of Estimated Parameters

	Canada				Mexico			
	RE		NRE		RE		NRE	
Parameters	Median	[5%, 95%]	Median	[5%, 95%]	Median	[5%, 95%]	Median	[5%, 95%]
θ^{-1}	NA	NA	7.56	[1.61 , 17.6]	NA	NA	504.2	[342.0. 733.9]
ω	3.53	[2.96, 4.07]	3.34	[2.74, 4.15]	5.15	[3.92, 6.92]	7.08	[6.36, 8.18]
φ	1.30	[0.41, 3.02]	0.70	[0.11, 2.07]	0.76	[0.17, 2.01]	1.10	[0.68, 2.07]
ρ_u	0.91	[0.84, 0.96]	0.89	[0.80, 0.95]	0.93	[0.86, 0.97]	0.08	[0.02, 0.19]
ρ_a	0.92	[0.85, 0.97]	0.88	[0.67, 0.96]	0.97	[0.94, 0.99]	0.70	[0.20, 0.92]
σ_u	0.029	[0.018, 0.050]	0.022	[0.015, 0.038]	0.036	[0.024, 0.060]	0.065	[0.053, 0.091]
σ_a	0.004	[0.001, 0.007]	0.008	[0.004, 0.012]	0.106	[0.081, 0.143]	0.019	[0.00, 0.035]

tional on the matrix of observables $\boldsymbol{O} \equiv [\pi_t, \Delta\tilde{e}_t]$. The posterior distribution, denoted as $P(\boldsymbol{\Theta}|\boldsymbol{O})$, is the product of the likelihood function of $\boldsymbol{O}$ and the prior distribution of $\boldsymbol{\Theta}$, which is denoted as $L(\boldsymbol{O}|\boldsymbol{\Theta})P(\boldsymbol{\Theta})$. The likelihood function $L(\boldsymbol{O}|\boldsymbol{\Theta})$ is evaluated numerically using the linear state space system (25)-(26) and Kalman filter. To evaluate the posterior distribution $P(\boldsymbol{\Theta}|\boldsymbol{O})$, we use a random-walk Metropolis-Hastings sampler described in Herbst and Schorfheide (2016). We use the last 1 million draws from the 5 million MCMC chains for the posterior evaluation with an acceptance rate of 25 percent. On a standard desktop computer[3], the estimation of the RE models takes less than five hours for each country. The estimation of the NRE models, on the other hand, is more computationally demanding, and requires approximately five days for each country.

4.5 Posterior Distributions: RE vs. NRE

Table 3 shows the posterior distributions of the parameters for Canada and Mexico, in both the RE and NRE models. In the case of Canada, the posterior median of θ^{-1} in the NRE model is 7.56, with a 90% credible interval of $[1.61, 17.6]$. The data tells us that the central bank has relatively small concerns about distorted expectations in Canada.

The slight departure of the distorted expectations of the NRE model from the RE in Canada results in the similarities of the posterior distributions of the other structural parameters between the two models. The posterior medians for ω in the RE and NRE models are 3.53 and 3.34, respectively, the posterior medians for φ in the two models are 1.30 and 0.70, respectively, and the 90% credible intervals for the two parameters have substantial overlaps between the two models. These results imply that the slopes for the aggregate supply and aggregate demand curves are similar between the two models. Using the posterior medians with the other calibrated parameters, the slopes of the aggregate demand curve (2), σ_α^{-1}, for the RE and NRE models are 2.01 and 1.94, respectively. The slope of the aggregate

[3]The estimations are performed on a desktop with a 3.60GHz Intel Core i7-7700 processor with four cores and 64.0 GB RAM.

supply curve (4), $\kappa \equiv \frac{(1-\zeta)(1-\zeta\beta)(\sigma_\alpha+\varphi)}{\zeta}$, for the RE and NRE models is 0.15 and 0.10, respectively. The small posterior median of θ^{-1} and the similarity of the equilibrium relation of aggregate demand and supply entail similar estimates for the exogenous processes. The posterior medians of the persistence of the cost-push shock ρ_u of the RE and NRE models are 0.91 and 0.89, respectively, and the posterior medians of the standard deviation of the cost-push shock σ_u for the two models are 0.029 and 0.022, respectively. The posterior medians of the persistence of the domestic productivity shock ρ_a for the RE and NRE models are 0.92 and 0.88, respectively, and the posterior medians of the standard deviation of the shock σ_a for the two models are 0.004 and 0.008, respectively. The 90% credible intervals for the parameters are similar across the two models. Accordingly, Canada shows slim differences between the estimated RE and NRE models overall.

In contrast, in the case of Mexico, there are non-trivial differences in the estimation results between the RE and NRE models. The posterior median of θ^{-1} in the NRE model is 504.2, with a 90% credible interval of $[342.0, 733.9]$. Based on the analyses in Section 3, we can infer that the Mexican data tells us that the central bank has substantial concerns about distorted expectations. The high θ^{-1} also affects the estimates of the other structural parameters. The posterior medians for ω and φ in the RE model are 5.15 and 0.76, respectively. However, the posterior medians for ω and φ in the NRE model are 7.08 and 1.10, respectively. The intersections of the 90% credible intervals of the same parameters between the two models are far narrower than in the case of Canada. The slopes of aggregate demand σ_α^{-1} for the two models become 2.03 and 2.51, and the slopes of aggregate supply κ in the RE and NRE models become 0.11 and 0.13 respectively. There are stark differences in the estimated parameters of the exogenous processes between the two models. For the parameters of the cost-push shock, the posterior medians of ρ_u and σ_u in the RE model are 0.93 and 0.036. In the NRE model, however, ρ_u and σ_u are 0.08 and 0.70, respectively. For the domestic productivity parameters, the posterior medians of ρ_a and σ_a in the RE model are 0.97 and 0.106, respectively, whereas they are 0.65 and 0.019, respectively, in the NRE model. In addition, there are few intersections between the 90% credible intervals for the same parameters of the two models. Accordingly, Mexico shows substantial differences between the estimated RE and NRE models overall[4].

[4]We also estimate NRE models for other small open economy countries: Australia, Chile, Colombia, New Zealand, Norway, South Africa, South Korea, and Sweden. We use observations of the countries starting in Q1:1991 to Q4:2008. We use the fitted AR (1) process of the U.S. real GDP and observations on the U.S. CPI inflation rate for the overlapping period to express the world output process and world inflation rate. Table C.8 in Appendix C summarizes posterior distributions of θ^{-1} in NRE models for these countries. Overall, developed countries (Australia, New Zealand, Norway, Sweden) have smaller θ^{-1} than emerging countries (Chile, Colombia, South Africa, South Korea), which is consistent with the difference in results between Canada and Mexico.

4.6 Prediction of Monetary Policy Rates: RE vs. NRE

Based on the estimated and calibrated parameters, the monetary policy rate in the model is computed as the solution in equation (17). How well the estimated models predict the actual monetary policy rate of the data is a question of interest. We compare the monetary policy rates predicted from the estimated models and in the data for Canada and Mexico. We use three-month interbank rates in Canada and immediate call rates in Mexico as proxies of actual monetary policy rates. All data are in quarterly frequency.

The left panel in (a) in Figure 9 compares actual monetary policy rate data[5] (black dotted line) and model-predicted policy rates from estimated RE (blue solid line) and NRE (red dashed line) models for Canada. The model-predicted policy rates are robustly optimal monetary policy rates generated by the models by using estimated parameters and the historical processes of shocks extracted from the Kalman smoother. The panel shows that there are few differences in the predictions of the monetary policy rate between the two models. In addition, the model-generated monetary policy rates show higher short-term volatilities than the actual data. Despite the differences in volatilities, the model-predicted interest rates and actual data show close dynamic comovements. To see the comovements more clearly, we compare the low frequency movements of the series which eliminate the short-term volatilities, as depicted in the right panel.[6] The model-generated monetary policy rates show strong comovements with the actual data. In terms of the distance from the actual data, the interest rate generated by the NRE model performs slightly better than that of the RE model. The small differences between the interest rates generated by the RE and NRE models are intuitive given that the posterior distributions of the two models are close and are associated with $\theta^{-1} = 7.56$, which is not far from zero.

The left panel in (b) in the same figure exhibits the three series for Mexico. Unlike Canada, the predictions of the monetary policy rates between the two models are sharply different in Mexico. More importantly, the NRE model obviously outperforms the RE model in the prediction. Throughout the sample period, the model-generated monetary policy rate by the NRE model almost replicates the path of the actual monetary policy rate. In particular, the policy rate from the NRE model succeeds in predicting the central bank's sharp interest rate hikes during the Mexican Peso Crisis. In contrast, the model-generated monetary policy rate by the RE model moves in the opposite direction from the actual policy rate throughout nearly all of the sample period. The right panel more clearly shows that the low frequency in the monetary policy rates between the data and the NRE model are very similar. In contrast, the RE model shows the opposite dynamics.

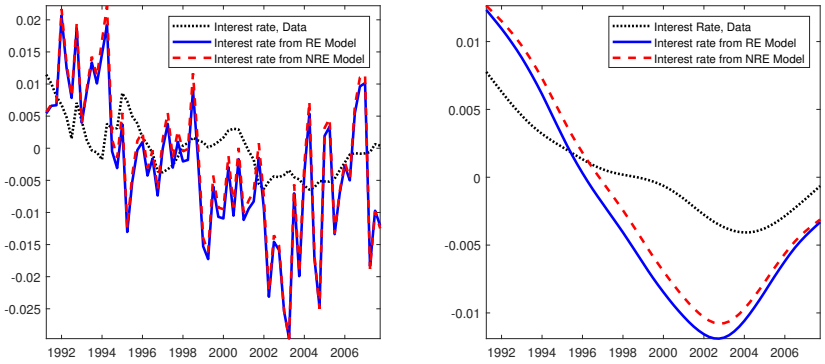
Given that no moments of the monetary policy rate are targeted in the estimation, the

[5] Since the original series are annualized rates, we quarterized and demeaned the series.

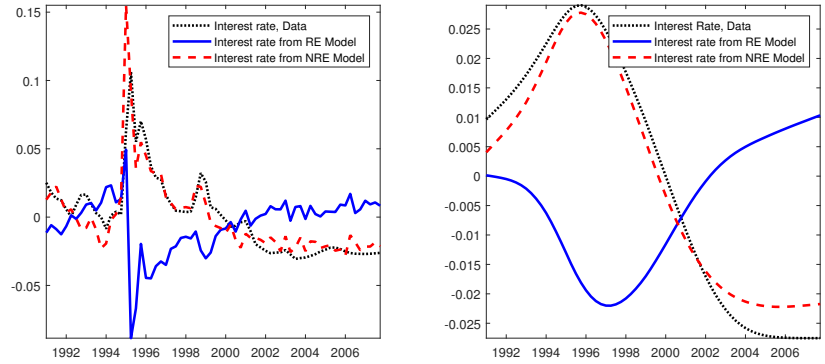
[6] We applied the Hodrick-Prescott filter with the penalty parameter 1600.

Figure 9: Monetary Policy Rates, Data and Model Predictions

(a) Canada



(b) Mexico



Notes. The left panels show the original three series, and the right panels show low-frequency components of the three series, extracted by HP filter.

finding that the monetary policy rate in the NRE model in Mexico closely matches that in the actual data is impressive. How is the finding related to θ^{-1} in the NRE model? Moreover, why is the performance of the RE model poor? Table 4 displays the second moments of the two observables used in the estimation, with the log marginal data density (MDD) of each model. The table provides insight that the persistence of inflation is a key identifying moment that distinguishes the NRE model from the RE benchmark.

In the case of Canada, the performance of the RE and NRE models in matching the moments of observations is qualitatively similar. The two models successfully match the standard deviations of observations and their cross-correlations. Both models are unsuccessful in matching high and positive autocorrelations of π_t for the first order, although the NRE model marginally improves the moments. The two models match the autocorrelations of $\Delta\tilde{e}_t$ well. In the case of Mexico, in contrast, there are obvious differences in the second moments

Table 4: Second Moments of Observables

	Canada			Mexico		
Moments	Data	RE	NRE	Data	RE	NRE
$\sigma(\pi_t)$	0.30	0.33	0.33	2.47	2.62	1.90
$\sigma(\Delta\tilde{e}_t)$	3.16	3.18	3.18	8.70	8.58	8.38
$\rho(\pi_t, \Delta\tilde{e}_t)$	-0.19	-0.25	-0.30	0.34	0.25	0.14
$\rho(\pi_t, \pi_{t-1})$	0.67	0.04	0.14	0.88	0.23	0.77
$\rho(\pi_t, \pi_{t-2})$	0.15	-0.01	0.04	0.67	0.09	0.44
$\rho(\pi_t, \pi_{t-4})$	-0.09	-0.04	-0.04	0.44	-0.02	0.12
$\rho(\Delta\tilde{e}_t, \Delta\tilde{e}_{t-1})$	0.30	0.35	0.38	0.08	0.11	-0.16
$\rho(\Delta\tilde{e}_t, \Delta\tilde{e}_{t-2})$	0.03	0.10	0.13	-0.02	0.04	-0.10
$\rho(\Delta\tilde{e}_t, \Delta\tilde{e}_{t-4})$	-0.07	-0.05	-0.05	0.00	-0.01	-0.03
Log MDD		436.34	417.67		213.31	263.40

Notes. The log marginal data densities are computed by using Geweke (1999)'s harmonic mean estimator.

between the two models. The RE model performs relatively better in matching the standard deviations of the observations and their cross-correlations, although the NRE model also matches them reasonably well. In terms of matching the autocorrelations of π_t, however, the NRE model remarkably outperforms the RE model. In the data, there is strong persistence of the CPI inflation rate up to higher orders. The first-, second-, and fourth-order autocorrelations are 0.88, 0.67, and 0.44, respectively, all of which are much higher than the moments in Canada. The RE model fails to match the moments. The RE model predicts 0.23, 0.09, and -0.02 for the first-, second-, and fourth-order autocorrelations, respectively. The NRE model predicts 0.77, 0.44, and 0.12 for the first-, second-, and fourth-order autocorrelations, respectively, which are much closer to the moments from the data. In terms of matching the autocorrelations of $\Delta\tilde{e}_t$, the two models both match them fairly well in the sense that the predicted autocorrelations are all close to zero, which are also shown in the data. The last row in the table shows that in the case of Canada, the MDD of the NRE model is smaller than that of the RE model. In the case of Mexico, the MDD of the NRE model increases by 25 percent relative to the RE model. This implies that the NRE model, which has one more estimable parameter θ^{-1}, does not necessarily always yield better performance, and the higher MDD of the NRE model for Mexico comes from its success in explaining inflation persistency.

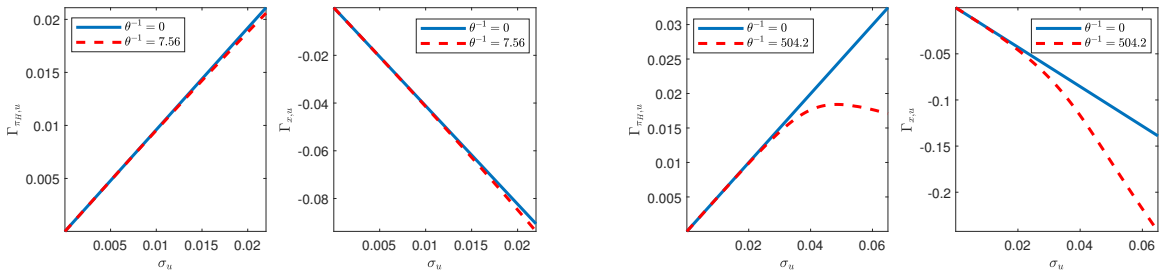
The critical insight on the success of the NRE model is in the conservative and more history-dependent monetary policy (than the RE benchmark) caused by the concerns about distorted expectations. Table 5 reports robustly optimal policy coefficients to cost push shock in the estimated models. In Mexico, the $\Gamma_{\pi_H,u}$ in the RE and NRE models are 0.0342

Table 5: Robustly Optimal Policy Coefficeints in Estimated Models

	Canada		Mexico	
	RE	NRE	RE	NRE
$\Gamma_{\pi_H,u}$	0.0209	0.0206	0.0342	0.0171
$\Gamma_{x,u}$	-0.0982	-0.0937	-0.1462	-0.2438

Figure 10: Policy Coefficients $\Gamma_{\pi_H,u}$ and $\Gamma_{x,u}$ in the NRE Models Conditional on σ_u

(a) Canada (b) Mexico

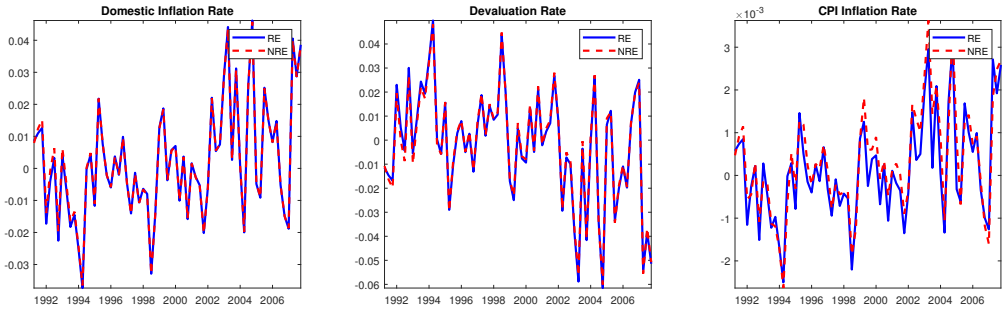


and 0.0171, respectively, and the $\Gamma_{x,u}$ in the two models are -0.1462 and -0.2438, respectively. The monetary policies in the two models have very different agendas, whereas the policies in Canada have commonality. The reason is because the posterior median $\theta^{-1} = 504.2$ of the NRE model in Mexico indeed produces a sluggish initial response of domestic inflation. Figure 10 shows the robustly policy coefficients of domestic inflation $\Gamma_{\pi_H,u}$ and the output gap $\Gamma_{x,u}$, conditional on the standard deviation σ_u. Except for σ_u and θ^{-1}, the parameters are set by using the calibrated and estimated parameters in the NRE model for the two countries. The maximum number of grids on the x-axis in each panel is the posterior median of σ_u of the NRE model for the two countries. The θ^{-1} for each country is the posterior median of the NRE model.

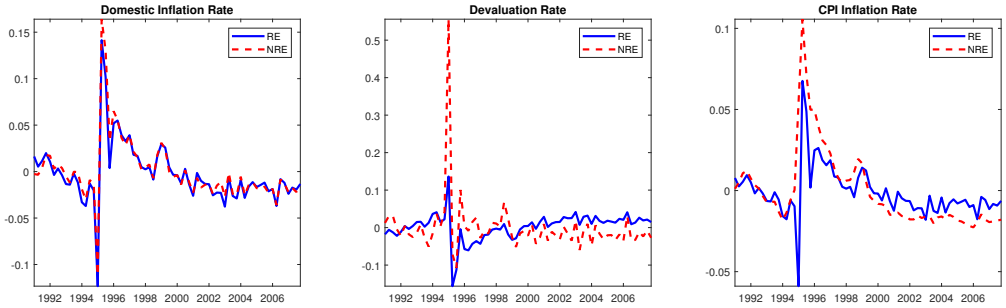
When $\theta^{-1} = 0$, $\Gamma_{\pi_H,u}$ increases linearly with the increase in σ_u in both countries, which clearly exhibits the certainty equivalence. In Canada (which is shown in the panels in (a)), there are slim differences between the policy coefficients in the RE and NRE models. In Mexico (which is shown in the panels in (b)), $\Gamma_{\pi_H,u}$ in the NRE model increases less than proportionally with σ_u. Furthermore, $\Gamma_{\pi_H,u}$ begins to decrease when σ_u is larger than approximately 0.04. The conservative monetary policy makes domestic inflation exhibit a sluggish response to the shock, which causes a dynamic persistence of domestic inflation. In contrast, the right panel shows that the $\Gamma_{x,u}$ of the NRE model declines more aggressively than in the case under RE as σ_u increases. The intuition is essentially the same as discussed in Section 3.

Figure 11: Historical Dynamics Model-Predicted Domestic Inflation, Devaluation, and CPI Inflation Driven by Cost-Push Shock

(a) Canada



(b) Mexico



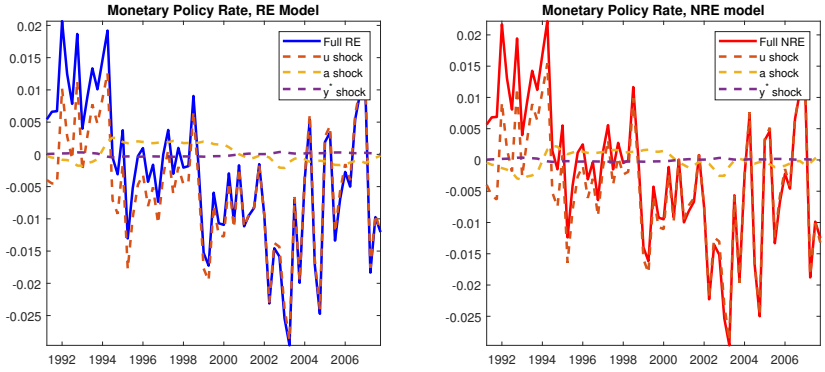
Notably, combining (7) and (23) gives the following relation between CPI inflation π_t, domestic inflation $\pi_{H,t}$, and the sum of the devaluation rate and the US inflation rate $\Delta\tilde{e}_t$:

$$\pi_t = (1-\alpha)\pi_{H,t} + \alpha\Delta\tilde{e}_t. \tag{27}$$

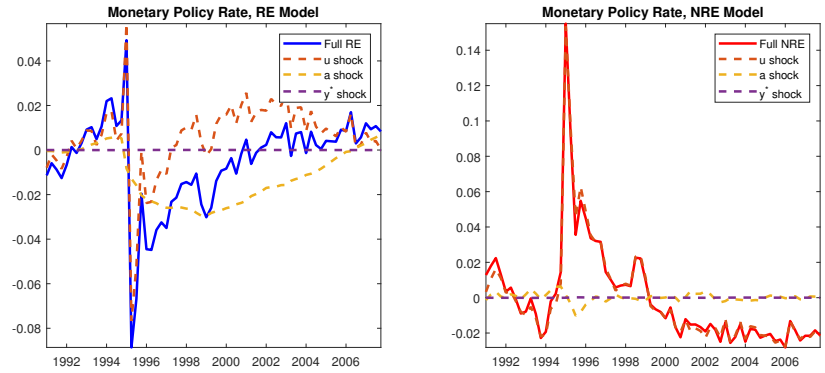
With the calibrated parameter $\alpha = 0.248$ in Mexico, (27) implies that the dynamics of CPI inflation have a high weight on the dynamics of domestic inflation $\pi_{H,t}$. However, the effect of the exchange rate pass-through from $\Delta\tilde{e}_t$ to π_t is also important, as is observed in Figure 8, particularly during the Mexican Peso Crisis, which involved large nominal devaluation. The estimated NRE model for Mexico can generate considerable responses of the devaluation rate subject to the cost-push shock, whereas the RE model has a limitation. The magnitude of policy coefficient $\Gamma_{x,u}$ in the NRE model is almost twice as large as that in the RE model. Remark that $\Delta\tilde{e}_t$ is an increasing function of Δx_t by equations (5)-(7). Thus, $\Gamma_{x,u}$ is tightly linked to the response of $\Delta\tilde{e}_t$. Figure 11 shows the model-predicted dynamics of $\pi_{H,t}$, Δe_t, and π_t, driven by historical cost-push shock produced by the Kalman smoother (see Figure C.23 in Appendix C for the historical dynamics of smoothed shocks). There are few differences in the dynamics of the three series between the RE and NRE models in Canada. In Mexico,

Figure 12: Historical Decomposition of Model-Predicted Monetary Policy Rates

(a) Canada



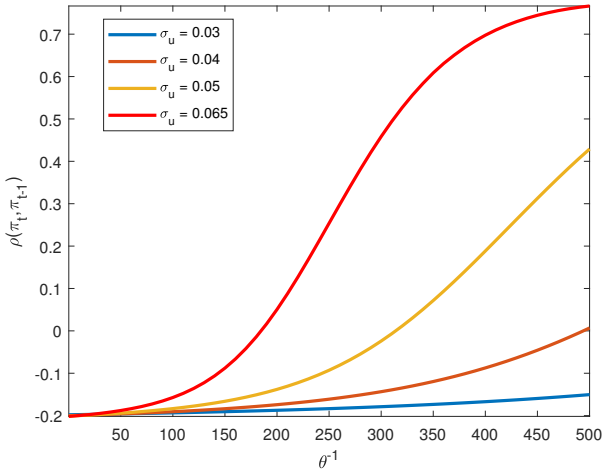
(b) Mexico



the dynamics of $\pi_{H,t}$ between the RE and NRE models are quite similar, and the NRE model generates a slightly more volatile movement during the Mexican Peso Crisis. Remark that in $\pi_{H,t}$ is determined by the sum of $\Phi_{\pi_H,t}$ and $\Gamma_{\pi_H,u}\varepsilon_t^u$. Thus, the dynamics of $\pi_{H,t}$ can be more volatile even though the optimal policy coefficient $\Gamma_{\pi_H,u}^{NRE}$ is less than $\Gamma_{\pi_H,u}^{RE}$ because of the impact from $\Phi_{\pi_H,t}^{NRE}$ and $\Phi_{\pi_H,t}^{RE}$ (see Figures C.26 - C.27 in Appendix C for decomposition of the variables driven by $\mathbf{\Phi}_t$ and $\mathbf{\Gamma}_u\varepsilon_t^u$). There are sharp differences in the dynamics of Δe_t. First, the NRE model predicts large devaluation during the Mexican Peso Crisis, whereas the devaluation predicted from the RE model for the same period is relatively limited. Second, the overall fluctuation beyond the Crisis periods is also substantially different. Because of the sharp differences in Δe_t, there are significant differences in the prediction of the CPI inflation π_t, which contains the pass-through effect of the devaluation.

The sharp differences observed in Figure 11 indicate that cost-push shock is not a primary driver of the fluctuations of devaluation in the RE model for Mexico, whereas it is in the NRE model. The differences in the prediction of the sources of the fluctuation lead to different predictions of the robustly optimal monetary policy rate. Figure 12 displays the historical

Figure 13: First-Order Autocorrelations of π_t, Conditional on θ^{-1} and σ_u in Mexico



decomposition of the model-predicted monetary policy rates. Unsurprisingly, there are few differences in the pattern of decompositions between the models in Canada, whereas there are sharp differences in Mexico. In Canada, the historical dynamics of the monetary policy rates in both models are responses to the cost-push shock (shown by the red dashed line). In contrast, the TFP shock plays a secondary role (shown by the yellow dashed line), and the world-output shock (shown by the purple dashed line) is negligible. In the RE model for Mexico, the TFP shock substantially affects the historical dynamics of policy rates. In the NRE model, on the other hand, the cost-push shock is the main driver of the historical responses of the policy rate, and the impacts of other shocks are negligible. In sum, Figures 11-12 imply that the estimated RE model in Mexico is misspecified because it assigns the TFP shock as the major source of the economic fluctuations, which leads to the failure in the prediction of the monetary policy rate consistent with the actual data.

Notably, the model-predicted CPI inflation driven by the cost-push shock in the NRE model for Mexico is also close to the actual observation of the CPI inflation rate. The result also indicates that the success of the NRE model of matching the persistency of the CPI inflation rate is closely related to the moments of the cost-push shock. Figure 13 shows the patterns of the first-order autocorrelation $corr(\pi_t, \pi_{t-1})$ of the NRE model for Mexico when θ^{-1} varies from 0 to 504.2 (posterior median in the NRE model), conditional on different σ_u from 0.03 to 0.065 (posterior median in the NRE model). We can observe that given any σ_u, $corr(\pi_t, \pi_{t-1})$ increases as θ^{-1} increases and further, $corr(\pi_t, \pi_{t-1})$ increases as σ_u becomes larger, conditional on any θ^{-1}. The interplay of the concerns of distorted expectations and robustly optimal monetary policy, and the volatility of the cost-push shock governs the inflation persistency in the NRE model.

Table 6: Posterior Distribution of θ^{-1} of NRE Models for Mexico: Subsamples

	Median	[5%, 95%]
Flexible Exchange Rate Regime	469.3	[284.0, 763.6]
Post-Peso Crisis	305.0	[182.0, 492.5]

4.7 Subsample Analysis

A potential concern on the estimation results for Mexico is that the sample periods of observables contain different exchange rate regimes - currency peg and the flexible exchange rate regime. Our benchmark results, including the high θ^{-1} might be mainly caused by the switch in the exchange rate regimes in the sample. Another potential concern is that the benchmark results could be altered when we exclude samples of the Mexican Peso Crisis.

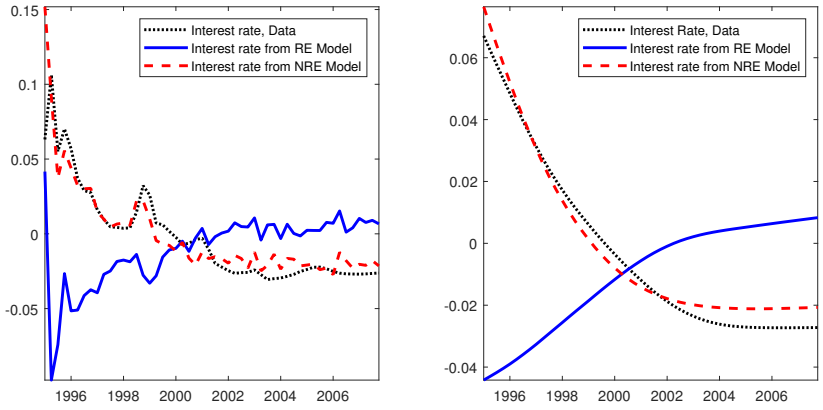
In this subsection, we argue that the impact of the potential concerns is limited. We rerun the estimations for Mexico using two subsamples: (i) a subsample covering the flexible exchange rate regime only (starting in 1Q:1995) and (ii) a subsample covering the period after the Mexican Peso Crisis (starting in 2Q:1995).

Table 6 summarizes the posterior distributions of θ^{-1} of the NRE models of Mexico. When considering the subsample covering the flexible exchange rate regime, the posterior median of θ^{-1} is 469.3, with a 90% credible interval [284.0, 763.6]. It is close to that in the benchmark estimation in Table 3. Thus, the benchmark result is not mainly caused by the switch in the exchange rate regimes in the sample. The result is intuitive because a high θ^{-1} is closely related to large fluctuations in the devaluation rate, which do not appear in the sample covering the currency peg period. Regarding the post-Peso Crisis subsample, the posterior median of θ^{-1} is 305.0 with a 90% credible interval [182.0, 492.5]. The distribution shows that the θ^{-1} becomes smaller than the benchmark. Thus, it also supports the claim that the large devaluations during the crisis favor a high estimate of θ^{-1}. However, note that Mexico continues to have much a higher θ^{-1} than Canada. This implies that the model predicts that the central bank in Mexico conducted robustly optimal monetary policy with substantial concerns about the private sector's NRE than the central bank in Canada, even when we focus on a subsample that excludes the Peso Crisis.

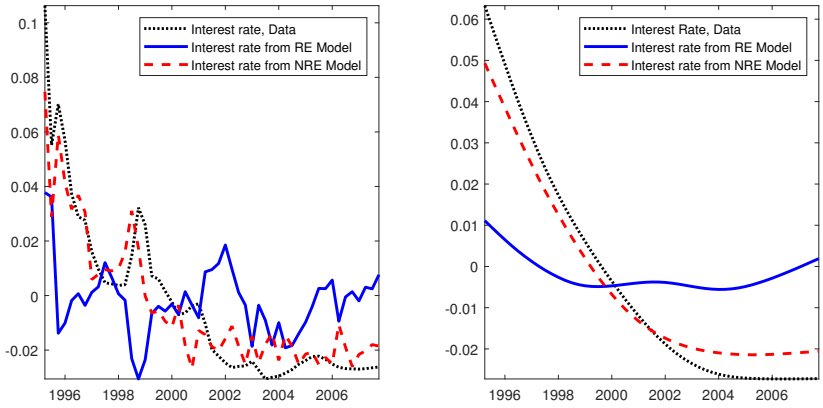
Moreover, an important theme of the estimation is to compare the performance of the prediction of monetary policy rates between the RE and NRE models. Figure 14 compares the performance of monetary policy predictions between the two models using the two different subsamples. The main finding in the benchmark estimation continues to hold: The NRE models succeed in predicting actual monetary policy rates, whereas the RE models fail to do so. In the subsample capturing the floating exchange rate regime, the monetary policy rates predicted by the NRE model are still very close to the actual data, and those predicted

Figure 14: Monetary Policy Rates, Data and Model Predictions for Mexico: Subsamples

(a) Flexible Exchange Rate Regime Subsample



(b) Post-Peso Crisis Subsample



Notes. The left panels show the original three series, and the right panels show low-frequency components of the three series, extracted by the HP filter.

by the RE model fails to match the direction. When using the post-Peso Crisis subsample, the NRE model continues to outperform the RE model.

5 Conclusion

In this paper, we study robustly optimal monetary policy in a small open economy where private agents' forward-looking expectations are potentially distorted. As the central bank's concerns about distorted expectations increase, it conducts a more history-dependent monetary policy than in the RE benchmark. When a cost-push shock occurs, the robustly optimal

policy yields a sluggish initial response and more history-dependent dynamics (than the RE benchmark) of domestic inflation and results in a greater initial response of the nominal devaluation rate. The estimated NRE models indicate small deviations from the RE benchmark in Canada, whereas there are substantial deviations in Mexico. The estimated NRE model for Mexico closely predicts the actual path of the monetary policy rate, whereas the estimated RE model fails to predict it. The key mechanism of the success of the NRE model for Mexico is that the robustly optimal monetary policy of the NRE model can explain key moments of the CPI inflation dynamics, which is a combination of domestic inflation dynamics and the exchange rate pass-through.

Considering models that move beyond the rational expectations hypothesis is increasingly important in macroeconomic modeling and policy analyses. The model we consider in this paper is parsimonious and accounts for an expectation wedge in the canonical NK environment. The model provides us with a meaningful improvement in the prediction of the actual monetary policy rate. We believe that considering richer model environments could better understand how distorted expectations relate to other shocks, wedges, and optimal policy designs. We leave such work to future research.

References

Adam, K. and M. Woodford (2012). Robustly optimal monetary policy in a microfounded New Keynesian model. *Journal of Monetary Economics 59*(5), 468–487.

Bhandari, A., J. Borovicka, and P. Ho (2019). Survey data and subjective beliefs in business cycle models. *working paper*.

Dennis, R. (2010). How robustness can lower the cost of discretion. *Journal of Monetary Economics 57*, 653–667.

Dennis, R., K. Leitemo, and F. Hammermann (2009). Monetary Policy in a Small Open Economy with a Preference for Robustness. *working paper*.

Gali, J. and T. Monacelli (2005). Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies 72*(3), 707–734.

Gerke, L. and F. Hammermann (2016). Robust Monetary Policy in a New Keynesian Model with Imperfect Interest Rate Pass-Through. *Macroeconomic Dynamics 20*, 1504–1526.

Geweke, J. (1999). Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Computation. *Econometric Reviews 18*(1), 1–126.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Gust, C., E. Herbst, and D. López-Salio (2020). Short-term Planning, Monetary Policy, and Macroeconomic Persistence. *Finance and Economics Discussion Series 2020-003. Washington: Board of Governors of the Federal Reserve System*.

Hansen, L. P. and T. J. Sargent (2001). Robust Control and Model Uncertainty. *American Economic Review 91*, 60–66.

Hansen, L. P. and T. J. Sargent (2008). *Robustness*. Princeton University Press.

Hansen, L. P. and T. J. Sargent (2012). Three types of ambiguity. *Journal of Monetary Economics 59*(5), 422–445.

Herbst, E. and F. Schorfheide (2016). *Bayesian Estimation of DSGE Models*. Princeton University Press.

Ilut, C. L. and M. Schneider (2014). Ambiguous Business Cycles. *American Economic Review 104*(8), 2368–2399.

Klein, P. (2000). Using the Generalized Shur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics and Control 24*, 1405–1423.

Kwon, H. and J. Miao (2017). Three types of robust Ramsey problems in a linear-quadratic framework. *Journal of Economic Dynamics and Control 76*, 211–231.

Kwon, H. and J. Miao (2019). Woodford's Approach to Robust Policy Analysis in a Linear-Quadratic Framework. *Macroeconomic Dynamics 23*(5), 1895–1920.

Leitemo, K. and U. Soderstrom (2008a). Robust Monetary Policy in a Small Open Economy. *Journal of Economic Dynamics and Control 32*(10), 3218–3252.

Leitemo, K. and U. Soderstrom (2008b). Robust Monetary Policy in the New Keynesian Framework. *Macroeconomic Dynamics 12*, 126–135.

Levine, P. and J. Pearlman (2010). Robust monetary rules under unstructured model uncertainty. *Journal of Economic Dynamics and Control 34*, 456–471.

Mendoza, E. (1991). Real Business Cycles in a Small-Open Economy. *American Economic Review 81*, 797–818.

Uribe, M. and S. Schmitt-Grohé (2017). *Open Economy Macroeconomics*. Princeton University Press.

Walsh, C. E. (2004). Robustly Optimal Instrument Rules and Robust Control: An Equivalence Result. *Journal of Money, Credit and Banking 36*(6), 1105–1113.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Woodford, M. (2010). Robustly Optimal Monetary Policy with Near-Rational Expectations. *American Economic Review 100*(1), 274–303.

Monetary Policy with Near-Rational Expectations in Open Economies

Yang Jiao Seunghoon Na Anurag Singh

Online Appendix

A Details of the Solution Method and the Proof of the Propositions

A.1 Solving the LQ System with Conditionally Linear Commitment using Methods in Kwon and Miao (2019)

We apply the solution method in Kwon and Miao (2019) to solve the linear-quadratic planner's problem with conditionally linear commitment. Solving the problem is a two-step process. In the first step, the hypothetical malevolent nature chooses m_{t+1} to maximize the welfare loss. The first-order condition of the Lagrangian with respect to m_{t+1} is:

$$\theta(1+\ln m_{t+1}) - \phi_t - \boldsymbol{\mu}'_{Y,t}\boldsymbol{Y}_{t+1} = 0,$$

which can be written in terms of $\boldsymbol{\varepsilon}_{t+1}$ by using the relation $\boldsymbol{y}_{t+1} = \boldsymbol{\Phi}_t + \boldsymbol{\Gamma}_t\boldsymbol{\varepsilon}_{t+1}, t \geq 0$. This relation along with constraint (1) yields the following expression:

$$m_{t+1} = \exp\left(-\frac{1}{2}\theta^{-2}\boldsymbol{\mu}'_{Y,t}\boldsymbol{\Gamma}_t\boldsymbol{\Gamma}'_t\boldsymbol{\mu}_{Y,t} + \theta^{-1}\boldsymbol{\mu}'_{Y,t}\boldsymbol{\Gamma}_t\boldsymbol{\varepsilon}_{t+1}\right). \tag{28}$$

Using the solution for the worst-case belief m_{t+1}, we can find the conditional expectations of $\boldsymbol{Y}_{t+1}$ and the relative entropy as:[7]

$$\begin{aligned}\mathbb{E}_t m_{t+1}\boldsymbol{Y}_{t+1} &= \boldsymbol{\Phi}_t + \theta^{-1}\boldsymbol{\Gamma}_t\boldsymbol{\Gamma}'_t\boldsymbol{\mu}_{Y,t},\\ \mathbb{E}_t m_{t+1}\ln m_{t+1} &= \frac{1}{2}\theta^{-2}\boldsymbol{\mu}'_{Y,t}\boldsymbol{\Gamma}_t\boldsymbol{\Gamma}'_t\boldsymbol{\mu}_{Y,t}.\end{aligned}$$

[7] First, we use the pdf under central bank's belief: $\ln f = -\frac{n_\varepsilon}{2}\ln(2\pi) - \frac{1}{2}\varepsilon'_{t+1}\varepsilon_{t+1}$. The probability density function under the worst-case belief can then be written as $\hat{f} = mf \implies \ln\hat{f} = \ln m + \ln f$, which results in a worst-case distribution that is normal with mean $\theta^{-1}\Gamma'_t\mu_{Y,t}$ and standard deviation I. Now, the conditional (on time-t) mean $\mathbb{E}_t m_{t+1}Y_{t+1} = E_t(m_{t+1}\Phi_t + m_{t+1}\Gamma_t\varepsilon_{t+1}) = \Phi_t + \mathbb{E}_t m_{t+1}\Gamma_t\varepsilon_{t+1} = \Phi_t + \hat{\mathbb{E}}_t\Gamma_t\varepsilon_{t+1} = \Phi_t + \Gamma_t\hat{\mathbb{E}}_t\varepsilon_{t+1} = \Phi_t + \theta^{-1}\Gamma_t\Gamma'_t\mu_{Y,t}$. Similarly, we can find $\mathbb{E}_t(m_{t+1}\ln m_{t+1})$.

In the second step of solving the problem, we can reconstruct the central bank's welfare loss function in matrix form as follows:

$$L(\boldsymbol{X}_t, \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t, i_t) = \frac{1}{2}\begin{bmatrix} \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t \end{bmatrix}' \boldsymbol{Q} \begin{bmatrix} \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t \end{bmatrix}$$

$$+\frac{1}{2} i_t \boldsymbol{R} i_t + \begin{bmatrix} \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t \end{bmatrix}' \boldsymbol{S} i_t,$$

where the matrix $\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}' & \boldsymbol{R} \end{bmatrix}$ is symmetric and positive definite.

The central bank chooses $\{\boldsymbol{X}_t, \boldsymbol{\Phi}_t, \boldsymbol{\Gamma}_t, i_t\}$ after substituting for the chosen value of m_{t+1} in the Lagrangian. The new Lagrangian is:

$$\mathcal{L} = \mathbb{E}_{-1} \sum_{t=0}^{\infty} \beta^t \Bigg\{ L(\boldsymbol{X}_t, \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t, i_t) - \frac{1}{2\theta}\boldsymbol{\mu}'_{Y,t}\boldsymbol{\Gamma}_t\boldsymbol{\Gamma}'_t\boldsymbol{\mu}_{Y,t} +$$

$$\begin{bmatrix} \boldsymbol{\mu}_{X,t+1} \\ \boldsymbol{\mu}_{Y,t} \end{bmatrix}' \left(\begin{bmatrix} \boldsymbol{X}_{t+1} \\ \boldsymbol{\Phi}_t + \frac{1}{\theta}\boldsymbol{\Gamma}_t\boldsymbol{\Gamma}'_t\boldsymbol{\mu}_{Y,t} \end{bmatrix} - \boldsymbol{A} \begin{bmatrix} \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t \end{bmatrix} - \boldsymbol{B} i_t \right) - \boldsymbol{C}\boldsymbol{\varepsilon}_{t+1} \Bigg\}. \tag{29}$$

The first-order necessary conditions with respect to $\{\boldsymbol{X}_t, \boldsymbol{\Phi}_t, \boldsymbol{\Gamma}_t, i_t\}$ are

$$\begin{aligned}
\mathbf{0} &= -\boldsymbol{Q}_{XX}\boldsymbol{X}_t - \boldsymbol{Q}_{XY}(\boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t) - \boldsymbol{S}_X i_t - \beta^{-1}\boldsymbol{\mu}_{X,t} + \boldsymbol{A}'_{XX}\mathbb{E}_t\boldsymbol{\mu}_{X,t+1} + \boldsymbol{A}'_{YX}\boldsymbol{\mu}_{Y,t}, \\
\mathbf{0} &= -\boldsymbol{R} i_t - \boldsymbol{S}'_X\boldsymbol{X}_t - \boldsymbol{S}'_Y(\boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t) + \boldsymbol{B}'_X\mathbb{E}_t\boldsymbol{\mu}_{X,t+1} + \boldsymbol{B}'_Y\boldsymbol{\mu}_{Y,t}, \\
\mathbf{0} &= -\left(\boldsymbol{Q}_{YY}\boldsymbol{\Phi}_t + \boldsymbol{Q}'_{XY}\mathbb{E}_t\boldsymbol{X}_{t+1}\right) - \boldsymbol{S}_Y\mathbb{E}_t i_{t+1} - \beta^{-1}\boldsymbol{\mu}_{Y,t} + \mathbb{E}_t\left[\boldsymbol{A}'_{XY}\boldsymbol{\mu}_{X,t+2} + \boldsymbol{A}'_{YY}\boldsymbol{\mu}_{Y,t+1}\right], \\
\mathbf{0} &= -\beta\boldsymbol{Q}_{YY}\mathbb{E}_t\left[\boldsymbol{\Phi}_t\boldsymbol{\varepsilon}'_{t+1} + \boldsymbol{\Gamma}_t\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}'_{t+1}\right] - \beta\boldsymbol{Q}'_{XY}\mathbb{E}_t\boldsymbol{X}_{t+1}\boldsymbol{\varepsilon}'_{t+1} \\
&\quad - \beta\boldsymbol{S}_Y\mathbb{E}_t i_{t+1}\boldsymbol{\varepsilon}'_{t+1} - \theta^{-1}\boldsymbol{\mu}_{Y,t}\boldsymbol{\mu}'_{Y,t}\boldsymbol{\Gamma}_t + \beta\mathbb{E}_t\left[\left(\boldsymbol{A}'_{XY}\boldsymbol{\mu}_{X,t+2} + \boldsymbol{A}'_{YY}\boldsymbol{\mu}_{Y,t+1}\right)\boldsymbol{\varepsilon}'_{t+1}\right]
\end{aligned} \tag{30}$$

where the matrices are partitioned as $\boldsymbol{A} \equiv \begin{bmatrix} \boldsymbol{A}_{XX} & \boldsymbol{A}_{XY} \\ \boldsymbol{A}_{YX} & \boldsymbol{A}_{YY} \end{bmatrix}$, $\boldsymbol{B} \equiv [\boldsymbol{B}_X, \boldsymbol{B}_Y]'$, and $\boldsymbol{C} \equiv [\boldsymbol{C}_X, \boldsymbol{C}_Y]'$, $\boldsymbol{Q} \equiv \begin{bmatrix} \boldsymbol{Q}_{XX} & \boldsymbol{Q}_{XY} \\ \boldsymbol{Q}'_{XY} & \boldsymbol{Q}_{YY} \end{bmatrix}$, and $\boldsymbol{S} \equiv \begin{bmatrix} \boldsymbol{S}_X \\ \boldsymbol{S}_Y \end{bmatrix}$. In addition, the first-order necessary

conditions with respect to $\{\boldsymbol{\mu}_{X,t+1}, \boldsymbol{\mu}_{Y,t}\}$ yield the law of motion of the constraints:

$$\begin{aligned}
\mathbf{0} &= \boldsymbol{X}_{t+1} - \boldsymbol{A}_{XX}\boldsymbol{X}_t - \boldsymbol{A}_{XY}(\boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t) - \boldsymbol{B}_X i_t - \boldsymbol{C}_X \boldsymbol{\varepsilon}_{t+1}, \\
\mathbf{0} &= \boldsymbol{\Phi}_t + \theta^{-1}\boldsymbol{\Gamma}_t\boldsymbol{\Gamma}_t'\boldsymbol{\mu}_{Y,t} - \boldsymbol{A}_{YX}\boldsymbol{X}_t - \boldsymbol{A}_{YY}(\boldsymbol{\Phi}_{t-1} + \boldsymbol{\Gamma}_{t-1}\boldsymbol{\varepsilon}_t) - \boldsymbol{B}_Y i_t,
\end{aligned}$$

Adding two more equations, $\mathbb{E}_t\boldsymbol{\varepsilon}_{t+1} = \mathbf{0}$ and $\mathbb{E}_t\boldsymbol{\mu}_{X,t+2} = \boldsymbol{E}_t\boldsymbol{\mu}_{X,t+1}$, to the aforementioned set of 5 equations, we solve:

$$\boldsymbol{J}\begin{bmatrix} \mathbb{E}_t\boldsymbol{\varepsilon}_{t+1} \\ \mathbb{E}_t\boldsymbol{X}_{t+1} \\ \boldsymbol{\Phi}_t \\ \mathbb{E}_t i_{t+1} \\ \mathbb{E}_t\boldsymbol{\mu}_{Y,t+1} \\ \mathbb{E}_t\boldsymbol{\mu}_{X,t+1} \\ \mathbb{E}_t\boldsymbol{\mu}_{X,t+2} \end{bmatrix} = \boldsymbol{F}\begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{X}_t \\ \boldsymbol{\Phi}_{t-1} \\ i_t \\ \boldsymbol{\mu}_{Y,t} \\ \boldsymbol{\mu}_{X,t} \\ \mathbb{E}_t\boldsymbol{\mu}_{X,t+1} \end{bmatrix} \tag{31}$$

where $\boldsymbol{\varepsilon}_t$, $\boldsymbol{X}_t$ and $\boldsymbol{\Phi}_{t-1}$ are predetermined state variables, whereas i_t, $\boldsymbol{\mu}_{Y,t}$, $\boldsymbol{\mu}_{X,t}$ and $\mathbb{E}_t\boldsymbol{\mu}_{X,t+1}$ are non-predetermined variables. The matrices $\boldsymbol{J}$ and $\boldsymbol{F}$ are

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{A}_{XX}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{B}_X' & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Q}_{XY}' & \boldsymbol{Q}_{YY} & \boldsymbol{S}_Y & -\boldsymbol{A}_{YY}' & \mathbf{0} & -\boldsymbol{A}_{XY}' \end{pmatrix},$$

$$\boldsymbol{F} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \boldsymbol{A}_{XY}\boldsymbol{\Gamma} & \boldsymbol{A}_{XX} & \boldsymbol{A}_{XY} & \boldsymbol{B}_X & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{A}_{YY}\boldsymbol{\Gamma} & \boldsymbol{A}_{YX} & \boldsymbol{A}_{YY} & \boldsymbol{B}_Y & -\theta^{-1}\boldsymbol{\Gamma}\boldsymbol{\Gamma}' & \mathbf{0} & \mathbf{0} \\ \boldsymbol{Q}_{XY}\boldsymbol{\Gamma} & \boldsymbol{Q}_{XX} & \boldsymbol{Q}_{XY} & \boldsymbol{S}_X & -\boldsymbol{A}'_{YX} & \beta^{-1}\boldsymbol{I} & \mathbf{0} \\ \boldsymbol{S}'_Y\boldsymbol{\Gamma} & \boldsymbol{S}'_X & \boldsymbol{S}'_Y & \boldsymbol{R} & -\boldsymbol{B}'_Y & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\beta^{-1}\boldsymbol{I} & \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

In our model, the partitions of the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{S}$ are produced below:

$$\boldsymbol{A}_{XX} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \rho_u & 0 & 0 \\ 0 & 0 & \rho_a & 0 \\ 0 & 0 & 0 & \rho_{y^*} \end{bmatrix}, \quad \boldsymbol{A}_{XY} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\boldsymbol{A}_{YX} = \begin{bmatrix} 0 & -\beta^{-1} & 0 & 0 \\ -\rho\sigma_\alpha^{-1} & \beta^{-1}\sigma_\alpha^{-1} & \Lambda_{y,a}(1-\rho_a) & -\sigma_\alpha^{-1}\varphi\Lambda_{y,y^*}(1-\rho_{y^*}) \end{bmatrix}, \quad \boldsymbol{A}_{YY} = \begin{bmatrix} \beta^{-1} & -\beta^{-1}\kappa_\alpha \\ -\beta^{-1}\sigma_\alpha^{-1} & 1+\beta^{-1}\sigma_\alpha^{-1}\kappa_\alpha \end{bmatrix}$$

$$\boldsymbol{B}_X = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{B}_Y = \begin{bmatrix} 0 \\ \sigma_\alpha^{-1} \end{bmatrix}, \quad \boldsymbol{C}_X = \begin{bmatrix} 0 & 0 & 0 \\ \sigma_u & 0 & 0 \\ 0 & \sigma_a & 0 \\ 0 & 0 & \sigma_{y^*} \end{bmatrix}, \quad \boldsymbol{C}_Y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{Q}_{XX} = \begin{bmatrix} \lambda_x(x^*)^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{Q}_{XY} = \begin{bmatrix} 0 & -\lambda_x x^* \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{Q}_{YX} = \boldsymbol{Q}'_{XY}, \quad \boldsymbol{Q}_{YY} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_x \end{bmatrix},$$

$$\boldsymbol{R} = [0], \quad \boldsymbol{S}_X = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{S}_Y = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We solve the system of linear difference equations (31) using Klein (2000)'s method. The solution to the aforementioned system therefore takes the state-space representation as in (16) and (17). After solving the system, taking the unconditional expectation on equation (30) and re-arranging it with respect to $\boldsymbol{\Gamma}_t$ (which should be constant in conditionally linear self-consistent commitment), we can obtain the following updating formula for $\boldsymbol{\Gamma}$:

$$\boldsymbol{\Gamma}^{(n)} = \beta \left(\theta^{-1}\mathbb{E}\left[\boldsymbol{\mu}_{Y,t}\boldsymbol{\mu}'_{Y,t}\right] + \beta \boldsymbol{Q}_{YY}\right)^{-1} \left\{\mathbb{E}[(\boldsymbol{A}'_{XY}\mu_{X,t+2} + \boldsymbol{A}'_{YY}\boldsymbol{\mu}_{Y,t+1})\boldsymbol{\varepsilon}'_{t+1}] - \boldsymbol{Q}_{XY}\boldsymbol{C}_X - \boldsymbol{S}_Y\mathbb{E}[i_t\boldsymbol{\varepsilon}'_t]\right\}, \tag{32}$$

where $\mathbb{E}[\boldsymbol{\mu}_{Y,t}\boldsymbol{\mu}'_{Y,t}], \mathbb{E}[\boldsymbol{\mu}_{X,t+2}\boldsymbol{\varepsilon}'_{t+1}], \mathbb{E}[\boldsymbol{\mu}_{Y,t+1}\boldsymbol{\varepsilon}'_{t+1}]$, and $\mathbb{E}[i_t\boldsymbol{\varepsilon}'_t]$ are computed as follows. First, define an auxiliary variable $\tilde{\mathbf{X}}_t \equiv [\boldsymbol{\varepsilon}'_t, \boldsymbol{X}'_t, \boldsymbol{\Phi}'_{t-1}]'$. Then from (16) we obtain

$$\mathbb{E}\left[\tilde{\boldsymbol{X}}_{t+1}\tilde{\boldsymbol{X}}'_{t+1}\right] = \boldsymbol{H}\left[\tilde{\boldsymbol{X}}_t\tilde{\boldsymbol{X}}'_t\right]\boldsymbol{H}' + \begin{pmatrix} \boldsymbol{I} & \boldsymbol{C}'_X & \boldsymbol{0} \\ \boldsymbol{C}_X & \boldsymbol{C}_X\boldsymbol{C}'_X & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

As long as the matrix $\boldsymbol{H}$ is stable, the above equation can be solved and we can obtain $\mathbb{E}\left[\tilde{\boldsymbol{X}}_{t+1}\tilde{\boldsymbol{X}}'_{t+1}\right]$. After that, we apply (17) and derive

$$\boldsymbol{\mu}_{Y,t} = \boldsymbol{G}_{2,\varepsilon}\boldsymbol{\varepsilon}_t + \boldsymbol{G}_{2,X}\boldsymbol{X}_t + \boldsymbol{G}_{2,\Phi}\boldsymbol{\Phi}_{t-1} = \boldsymbol{N}_2\tilde{\boldsymbol{X}}_t,$$

where $\boldsymbol{G} \equiv [\boldsymbol{G}_{2,\varepsilon}, \boldsymbol{G}_{2,X}, \boldsymbol{G}_{2,\Phi}]$ is the second row of $\boldsymbol{G}$ in (17). Then we obtain

$$\mathbb{E}\left[\boldsymbol{\mu}_{Y,t}\boldsymbol{\mu}'_{Y,t}\right] = \boldsymbol{G}_2\mathbb{E}\left[\tilde{\boldsymbol{X}}_t\tilde{\boldsymbol{X}}'_t\right]\boldsymbol{G}_2,$$

$$\mathbb{E}\begin{pmatrix} i_{t+1}\boldsymbol{\varepsilon}'_{t+1} \\ \boldsymbol{\mu}_{Y,t+1}\boldsymbol{\varepsilon}'_{t+1} \\ \boldsymbol{\mu}_{X,t+1}\boldsymbol{\varepsilon}'_{t+1} \end{pmatrix} = \boldsymbol{G}\mathbb{E}\begin{pmatrix} \boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}'_{t+1} \\ \boldsymbol{X}_{t+1}\boldsymbol{\varepsilon}'_{t+1} \\ \boldsymbol{\Phi}_t\boldsymbol{\varepsilon}'_{t+1} \end{pmatrix} = \boldsymbol{G}\begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{C}_X \\ \boldsymbol{0} \end{pmatrix}.$$

The updating procedure of $\boldsymbol{\Gamma}$ in (32) stops when $||\boldsymbol{\Gamma}^{(n-1)} - \boldsymbol{\Gamma}^{(n)}|| < \xi$. We set $\xi = 10^{-5}$.

A.2 Proof of the Propositions

Proof of Proposition 3.1. The interest rate i_t is a choice variable of the central bank to determine the optimal path of domestic inflation $\pi_{H,t}$ and output gap x_t. At any given equilibrium path of $\{\pi_{H,t}, x_t\}$ and worst-case expectations, the corresponding path of interest rate makes the aggregate demand equation (2) always hold. Thus, the path of $\{\pi_{H,t}, x_t\}$ can be determined without reference to equation (2), as long as the zero lower bound is not binding, which is implicitly assumed in the model. Then the natural rate shocks $\{a_t, y_t^*\}$ in (2) do not affect the optimal paths of $\{\pi_{H,t}, x_t\}$. Thus, $\Gamma_{\pi_H,a} = \Gamma_{x,a} = \Gamma_{\pi_H,y^*} = \Gamma_{x,y^*} = 0$. □

Proof of Proposition 3.2. In the equilibrium with the robustly optimal monetary policy, the expectations of $\boldsymbol{Y}_{t+1}$ conditional on the worst-case beliefs are determined as (19). Then, proposition 3.1 implies that $\mu_{x,t} = 0,\ \forall t$, i.e., the component in $\boldsymbol{\mu}_{Y,t}$ regarding (2) is zero. Thus, for any $\theta \in \mathbb{R}_+$, $\hat{\mathbb{E}}_t x_{t+1} \equiv \mathbb{E}_t m_{t+1} x_{t+1} = \Phi_{x,t} + \theta^{-1}\Gamma_{x,t}^2 \mu_{x,t} = \Phi_{x,t} = \mathbb{E}_t x_{t+1}$. □

Proof of Proposition 3.3. Given the worst-case NRE from private agents, the first-order condition of the minimization problem of (29) with respect to $\Phi_{\pi_H,t}$ is summarized as follows (we follow the method in Appendix A.2 in Woodford (2010)):

$$-\beta\frac{\lambda_x}{\kappa}\frac{\pi_{H,t} - u_t - \kappa\bar{x} - \beta\Phi_{\pi_H,t}}{\Xi} + \beta\mathbb{E}_t\left(\pi_{H,t+1} + \frac{\lambda_x}{\kappa^2}\frac{\pi_{H,t+1} - u_{t+1} - \kappa\bar{x} - \beta\Phi_{\pi_H,t+1}}{\Xi}\right) = 0$$

where

$$\Xi = 1 - \theta^{-1}\beta^2\frac{\lambda_x}{\kappa^2}\Gamma_{\pi_H,u}^2 > 0.$$

Under the restriction of the linear commitment policy, we have

$$\beta\frac{\lambda_x}{\kappa}\frac{\Phi_{\pi_H,t-1} - \Gamma_{\pi_H,u}\varepsilon_t^u - (\rho_u u_{t-1} + \sigma_u\varepsilon_t^u) - \beta\Phi_{\pi_H,t}}{\Xi} + \beta\left(\Phi_{\pi_H,t} + \frac{\lambda_x}{\kappa^2}\frac{\Phi_{\pi_H,t} - \rho_u u_t - \beta\mathbb{E}_t\Phi_{\pi_H,t+1}}{\Xi}\right), \tag{33}$$

Then, (33) is rewritten as the second-order stochastic difference equation of the process $\{\Phi_{\pi,u,t}\}$ as follows:

$$\mathbb{E}_t\left[A(L)\Phi_{\pi_H,u,t+1}\right] = (\sigma_u - \Gamma_{\pi_H,u})\varepsilon_t^u - \rho_u(u_t - u_{t-1}),$$

where

$$A(L) \equiv \beta - \left(1 + \beta + \frac{\kappa^2\Xi}{\lambda_x}\right)L + L^2, \tag{34}$$

where L is the lag operator. By factorizing the lag polynomial (34), we have the unique and

stationary solution for $\Phi_{\pi,t}$ as follows:

$$\Phi_{\pi_H,t} = \mu\Phi_{\pi_H,t-1} - \mu\left((\sigma_u - \Gamma_{\pi_H,u})\varepsilon_t^u - \rho_u(u_t - u_{t-1})\right), \tag{35}$$

where $0 < \mu < 1$ is the smaller root of the characteristic equation (we follow the method in Appendix C in Woodford (2003)),

$$\mathbb{P}(\mu) \equiv \beta\mu^2 - \left(1 + \beta + \frac{\kappa^2\Xi}{\lambda_x}\right)\mu + 1 = 0.$$

Finally, setting $\rho_{\pi_H} \equiv \mu$ and applying the relation of optimal commitment $\pi_{H,t} = \Phi_{\pi_H,t-1} + \Gamma_{\pi_H,u}\varepsilon_t^u$ to (34) gives equation (22). □

B Description of The Linear State Space System for Estimation

The linear state space system, governing the solution to the equilibrium is given as:

$$\boldsymbol{\xi}_{t+1} = \boldsymbol{H}\boldsymbol{\xi}_t + \boldsymbol{S}^{\varepsilon}\boldsymbol{\varepsilon}_{t+1},$$

$$\boldsymbol{\chi}_t = \boldsymbol{G}\boldsymbol{\xi}_t,$$

where $\boldsymbol{\xi}_t \equiv [\epsilon_t^u, \epsilon_t^a, \epsilon_t^y, 1, u_t, a_t, y_t^*, \Phi_{\pi_H,t-1}, \Phi_{x,t-1}]'$, $\boldsymbol{\chi}_t \equiv [i_t, \mu_{\pi_H,t}, \mu_{x,t}, \mu_{1,t}, \mu_{u,t}, \mu_{a,t}, \mu_{y^*,t}, \mu_{1,t+1}, \mu_{u,t+1}, \mu_{a,t+}$ $\mu_{y^*,t+1}]'$, $\boldsymbol{S}^{\varepsilon} \equiv [\boldsymbol{I}, \boldsymbol{C}_X', \boldsymbol{0}]'$, and $\boldsymbol{\varepsilon}_{t+1} \equiv [\epsilon_t^u, \epsilon_t^a, \epsilon_t^y]$. The state variables in this system are used to construct the law of motion of the observables in the estimation process.

There are two variables that are used as observables in the estimation. The CPI inflation, π_t, is the first observable. In the model, it is determined as:

$$\begin{aligned}
\pi_t &= \pi_{H,t} + \alpha\Delta s_t \\
&= \left(\Gamma_{\pi_H,u,t-1} + \alpha\sigma_\alpha\Gamma_{x,u,t-1}\right)\epsilon_t^u - \alpha\sigma_\alpha\Gamma_{x,u,t-2}\epsilon_{t-1}^u + \left(\Gamma_{\pi_H,a,t-1} + \alpha\sigma_\alpha\Gamma_{x,a,t-1}\right)\epsilon_t^a - \alpha\sigma_\alpha\Gamma_{x,a,t-2}\epsilon_{t-1}^a \\
&+ \left(\Gamma_{\pi_H,y^*,t-1} + \alpha\sigma_\alpha\Gamma_{x,y^*,t-1}\right)\epsilon_t^{y^*} - \alpha\sigma_\alpha\Gamma_{x,y^*,t-2}\epsilon_{t-1}^{y^*} + \Phi_{\pi_H,t-1} + \alpha\sigma_\alpha\Phi_{x,t-1} - \alpha\sigma_\alpha\Phi_{x,t-2} \\
&+ \alpha\sigma_\alpha\left[\frac{1+\varphi}{\sigma_\alpha+\varphi}(a_t - a_{t-1}) - \frac{\sigma_\alpha+\varphi+\alpha(\omega-1)\sigma_\alpha}{\sigma_\alpha+\varphi}(y_t^* - y_{t-1}^*)\right]
\end{aligned}$$

The second observable is the devaluation of the domestic economy after adjusting for US CPI inflation, $\Delta\tilde{e}_t$. In the model, it is determined as:

$$\begin{aligned}
\Delta\tilde{e}_t &= \Delta e_t + \pi_t^* \\
&= \pi_{H,t} + \Delta s_t \\
&= \left(\Gamma_{\pi_H,u,t-1} + \sigma_\alpha\Gamma_{x,u,t-1}\right)\epsilon_t^u - \sigma_\alpha\Gamma_{x,u,t-2}\epsilon_{t-1}^u + \left(\Gamma_{\pi_H,a,t-1} + \sigma_\alpha\Gamma_{x,a,t-1}\right)\epsilon_t^a - \sigma_\alpha\Gamma_{x,a,t-2}\epsilon_{t-1}^a \\
&+ \left(\Gamma_{\pi_H,y^*,t-1} + \sigma_\alpha\Gamma_{x,y^*,t-1}\right)\epsilon_t^{y^*} - \sigma_\alpha\Gamma_{x,y^*,t-2}\epsilon_{t-1}^{y^*} + \Phi_{\pi_H,t-1} + \sigma_\alpha\Phi_{x,t-1} - \sigma_\alpha\Phi_{x,t-2} \\
&+ \sigma_\alpha\left[\frac{1+\varphi}{\sigma_\alpha+\varphi}(a_t - a_{t-1}) - \frac{\sigma_\alpha+\varphi+\alpha(\omega-1)\sigma_\alpha}{\sigma_\alpha+\varphi}(y_t^* - y_{t-1}^*)\right]
\end{aligned}$$

The two observables, therefore, constitute the measurement equation, which can be written in terms of the state variables from the equilibrium with the robustly optimal monetary policy, $\boldsymbol{\xi}_t$, and their lagged counterpart, $\boldsymbol{\xi}_{t-1}$. After accounting for measurement errors, the

measurement equation is given as:

$$\begin{bmatrix} \pi_t \\ \Delta\tilde{e}_t \end{bmatrix} = \tilde{\boldsymbol{G}} \cdot \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_{t-1} \end{bmatrix} + \begin{bmatrix} \sigma_\pi^{me} & 0 \\ 0 & \sigma_{\Delta\tilde{e}}^{me} \end{bmatrix} \cdot \begin{bmatrix} \epsilon_t^{me,\pi} \\ \epsilon_t^{me,\Delta\tilde{e}} \end{bmatrix}$$

The corresponding transition equation, to be used in the estimation, is given as:

$$\begin{bmatrix} \boldsymbol{\xi}_{t+1} \\ \boldsymbol{\xi}_t \end{bmatrix} = \tilde{\boldsymbol{H}} \cdot \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_{t-1} \end{bmatrix} + \boldsymbol{\nu}_{t+1}$$

where the matrices $\tilde{\boldsymbol{G}}$ and $\tilde{\boldsymbol{H}}$ are given as:

$$\tilde{\boldsymbol{G}} = \begin{bmatrix} \Gamma_{\pi_H,u} + \alpha\sigma_\alpha\Gamma_{x,u} & \Gamma_{\pi_H,u} + \sigma_\alpha\Gamma_{x,u} \\ \Gamma_{\pi_H,a} + \alpha\sigma_\alpha\Gamma_{x,a} & \Gamma_{\pi_H,a} + \sigma_\alpha\Gamma_{x,a} \\ \Gamma_{\pi_H,y^*} + \alpha\sigma_\alpha\Gamma_{x,y^*} & \Gamma_{\pi_H,y^*} + \sigma_\alpha\Gamma_{x,y^*} \\ 0 & 0 \\ 0 & 0 \\ \alpha\sigma_\alpha\frac{1+\varphi}{\sigma_\alpha+\varphi} & \sigma_\alpha\frac{1+\varphi}{\sigma_\alpha+\varphi} \\ -\alpha\sigma_\alpha\frac{\sigma_\alpha+\varphi+(\omega-1)\alpha\sigma_\alpha}{\sigma_\alpha+\varphi} & -\sigma_\alpha\frac{\sigma_\alpha+\varphi+(\omega-1)\alpha\sigma_\alpha}{\sigma_\alpha+\varphi} \\ 1 & 1 \\ \alpha\sigma_\alpha & \sigma_\alpha \\ -\alpha\sigma_\alpha\Gamma_{x,u} & -\sigma_\alpha\Gamma_{x,u} \\ -\alpha\sigma_\alpha\Gamma_{x,a} & -\sigma_\alpha\Gamma_{x,a} \\ -\alpha\sigma_\alpha\Gamma_{x,y^*} & -\sigma_\alpha\Gamma_{x,y^*} \\ 0 & 0 \\ 0 & 0 \\ -\alpha\sigma_\alpha\frac{1+\varphi}{\sigma_\alpha+\varphi} & -\sigma_\alpha\frac{1+\varphi}{\sigma_\alpha+\varphi} \\ \alpha\sigma_\alpha\frac{\sigma_\alpha+\varphi+(\omega-1)\alpha\sigma_\alpha}{\sigma_\alpha+\varphi} & \sigma_\alpha\frac{\sigma_\alpha+\varphi+(\omega-1)\alpha\sigma_\alpha}{\sigma_\alpha+\varphi} \\ 0 & 0 \\ -\alpha\sigma_\alpha & -\sigma_\alpha \end{bmatrix}^T,$$

$$\tilde{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{H} & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} \end{bmatrix},$$

and the vector $\boldsymbol{\nu}_{t+1}$ is:

$$\boldsymbol{\nu}_{t+1} = \begin{bmatrix} \boldsymbol{S}^{\varepsilon} \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \epsilon^{u}_{t+1} \\ \epsilon^{a}_{t+1} \\ \epsilon^{y}_{t+1} \end{bmatrix}.$$

C Supplementary Tables and Figures

Table C.7: Parameters

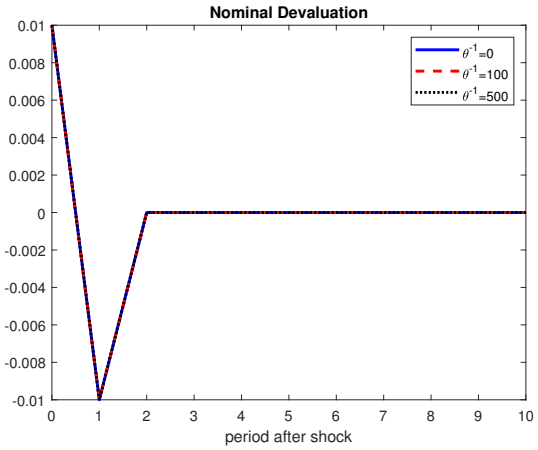
Parameter	Definition of the parameter	Parameter value
(1) Parameters that govern the distorted SOE-NK equilibrium dynamics		
(1.1) Primary structural parameters		
β	Discount factor	C-Specific (0.99 for CAN)
ρ	Governs the natural rate of interest, $\rho = \beta^{-1} - 1$	C-Specific (0.01 for CAN)
σ	Governs the curvature of the utility function	1
α	Governs the degree of home bias or openness of the economy	C-Specific (0.20 for CAN)
η	Measures the substitutability between domestic and foreign goods	NA
ϵ	Measures the substitutability between different home varieties	6
γ	Measures the substitutability between goods produced in different foreign countries	NA
ζ	Calvo-Yun parameter	0.75
φ	Governs the Frisch elasticity of labor supply	C-Specific (Estimated)
(1.2) Secondary structural parameters		
ω	Governs the effect of changes in the terms of trade on output, $\omega = \sigma\gamma + (1-\alpha)(\sigma\eta - 1)$	C-Specific (Estimated)
Θ	Related to ω, $\Theta = (\omega - 1)$	C-Specific (Comes from ω)
σ_α	Captures the same thing as ω, $\sigma_\alpha = \sigma/(1-\alpha+\alpha\omega)$	C-Specific (Comes from ω and α)
λ	Governs the parameter that defines the slope of the Phillips curve, $\lambda = (1-\beta\theta)(1-\theta)/\theta$	C-Specific (0.0858 for CAN)
κ_α	Slope of Phillips curve, $\kappa_\alpha = \lambda(\sigma_\alpha + \varphi)$	C-Specific (Comes from ω, λ and φ)
$\Lambda_{y,0}$	Intercept on the natural level of output	0
$\Lambda_{y,a}$	Governs the effect of the domestic TFP on the natural level of output: $\frac{1+\varphi}{\sigma_\alpha+\varphi}$	C-Specific (Comes from ω and φ)
Λ_{y,y^*}	Governs the effect of the world output on the natural level of output: $-\frac{\alpha\Theta\sigma_\alpha}{\sigma_\alpha+\varphi}$	C-Specific (Comes from ω and φ)
$\Lambda_{r,a}$	Governs the effect of the domestic TFP on the natural interest rate: $-\sigma_\alpha\Lambda_{y,a}(1-\rho_a)$	C-Specific (Comes from ω and φ)
Λ_{r,y^*}	Governs the effect of the world output on the natural interest rate: $-\varphi\Lambda_{y,y^*}(1-\rho_y^*)$	C-Specific (Comes from ω and φ)
(1.3) Parameters that govern the shock processes		
ρ_u	Persistence of cost-push shock process	C-Specific (Estimated)
ρ_a	Persistence of TFP shock process	C-Specific (Estimated)
ρ_{y^*}	Persistence of world-output shock process	0.9261 (Using US output process)
σ_u	Standard deviation of cost-push shock process	C-Specific (Estimated)
σ_a	Standard deviation of TFP shock process	C-Specific (Estimated)
σ_{y^*}	Standard deviation of world-output shock process	0.0055 (Using US output process)
(2) Parameters that govern the central bank's problem		
λ_x	Governs the relative weight that the policymaker assigns to output stabilization, $\lambda_x = (1+\varphi)\lambda/\epsilon$	C-Specific (Comes from ϵ, λ and φ)
$\bar{x}$	Governs the output gap target	0
θ	Governs the robustness concern of the policymaker	C-Specific (Estimated)
(3) Parameters that govern measurement error		
σ_π^{me}	Standard deviation in the measurement error of CPI inflation	C-Specific (Estimated)
$\sigma_{\Delta e}^{me}$	Standard deviation in the measurement error of nominal devaluation plus world inflation	C-Specific (Estimated)

Table C.8: Posterior Distributions for θ^{-1}, Other Countries

	θ^{-1}, Median	[5%, 95%]
Australia	67.9	[8.6, 179.2]
Chile	279.0	[160.9, 476.0]
Colombia	451.2	[225.0, 865.4]
New Zealand	9.2	[0.8, 56.5]
Norway	11.0	[0.9, 35.6]
South Africa	363.7	[241.1, 542.5]
South Korea	94.1	[27.6, 173.2]
Sweden	17.7	[1.19, 151.2]

Figure C.15: Dynamic Responses of Nominal Devaluation Rate to Domestic Productivity Shock

(a) $\rho_a = 0$



(b) $\rho_a = 0.9$

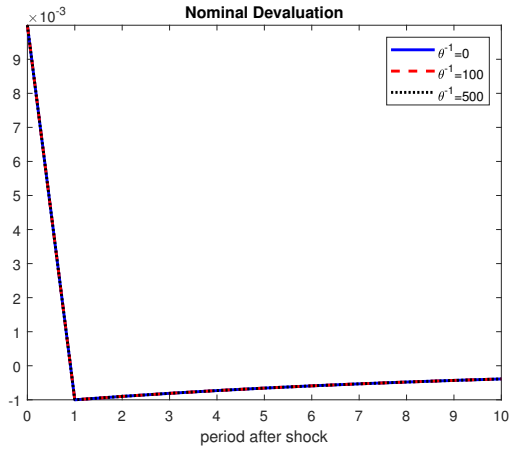
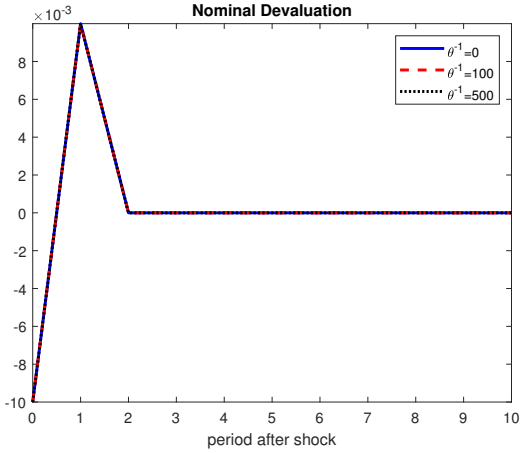


Figure C.16: Dynamic Responses of Nominal Devaluation Rate to World Output Shock

(a) $\rho_{y^*} = 0$



(b) $\rho_{y^*} = 0.9$

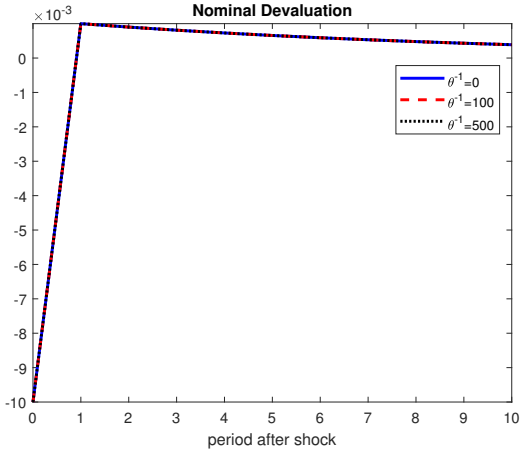


Figure C.17: Log Posteriors, Accepted Chains

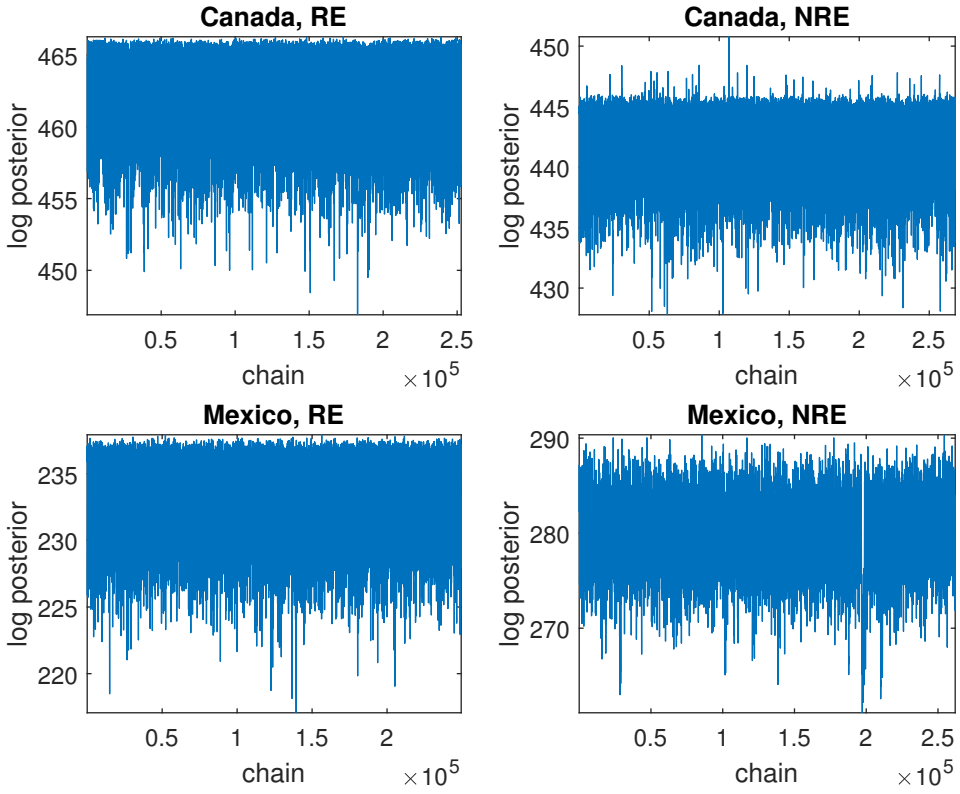


Figure C.18: Prior (red) and Posterior Distributions of the Structural Parameters in the RE Model, Canada

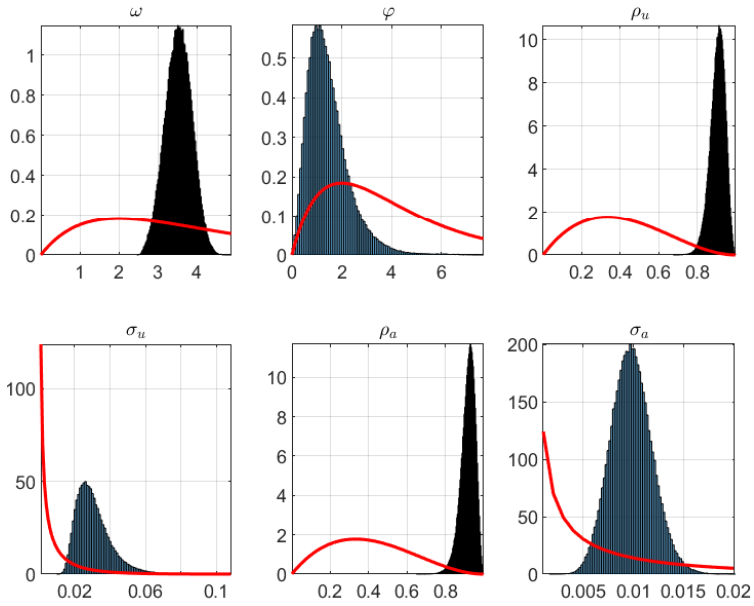


Figure C.19: Prior (red) and Posterior Distributions of the Structural Parameters in the RE Model, Mexico

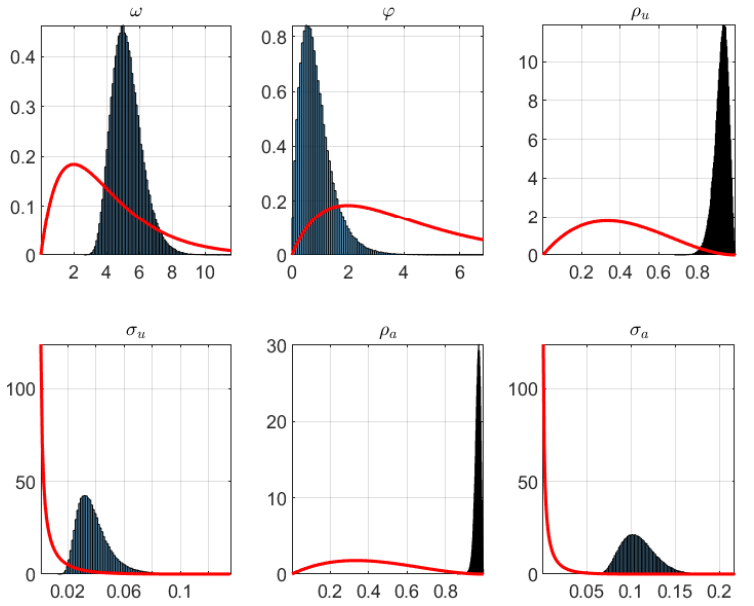


Figure C.20: Prior (red) and Posterior Distributions of θ^{-1} in the NRE Model, Canada

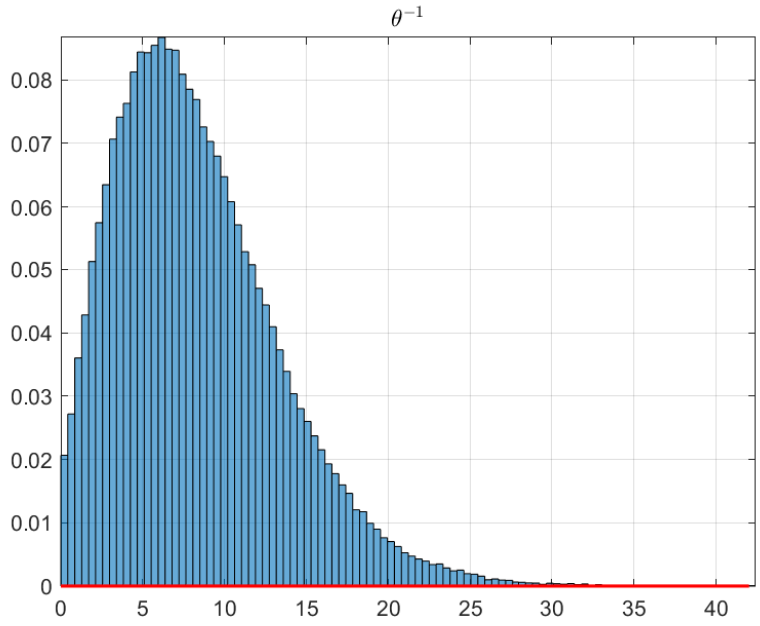


Figure C.21: Prior (red) and Posterior Distributions of Other Structural Parameters in the NRE Model, Canada

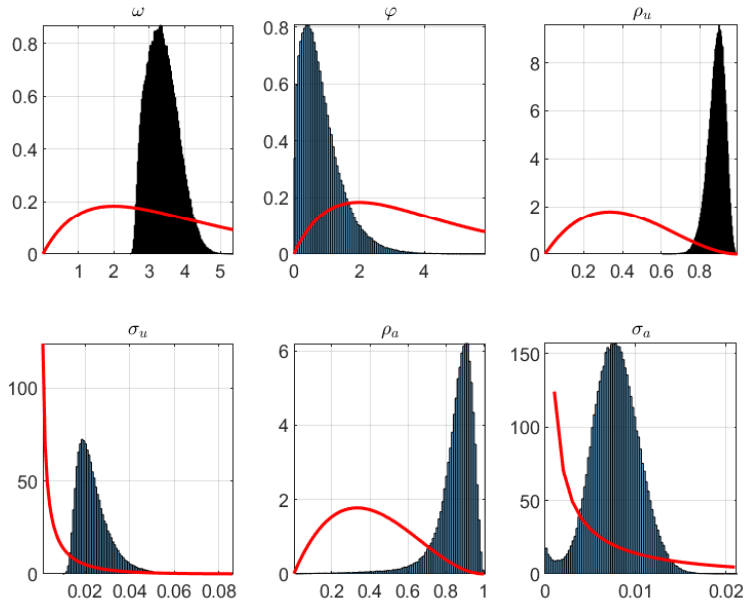


Figure C.22: Prior (red) and Posterior Distributions of θ^{-1} in the NRE Model, Mexico

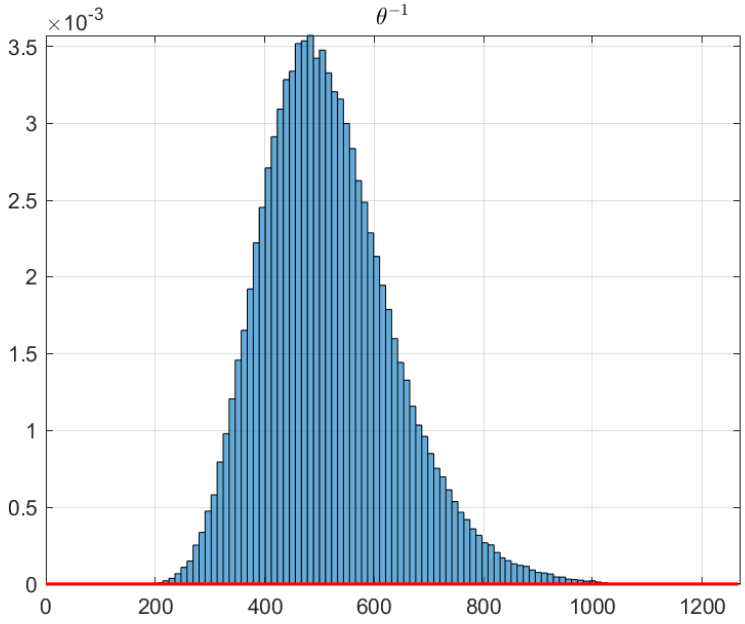


Figure C.23: Prior (red) and Posterior Distributions of Other Structural Parameters in the NRE Model, Mexico

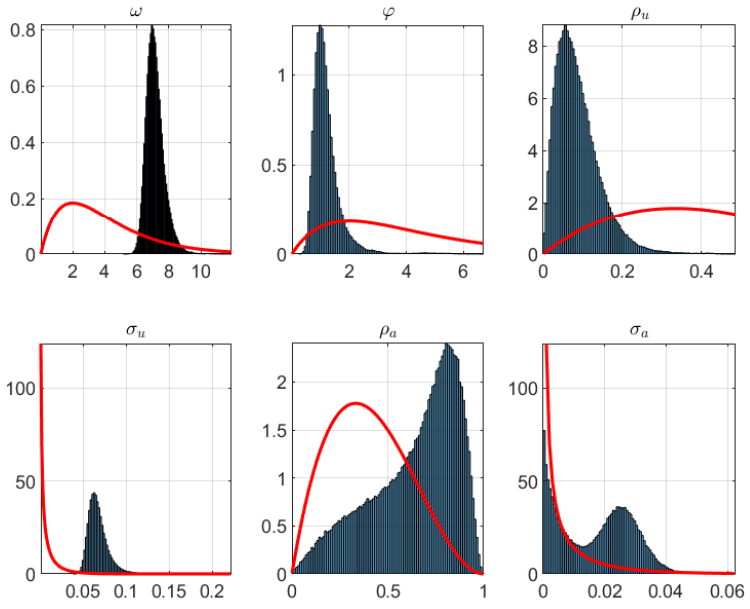
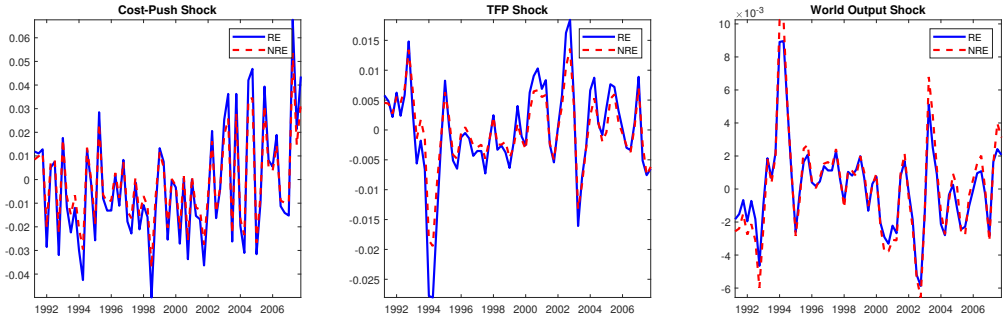


Figure C.24: Smoothed Shocks in the Estimated Models

(a) Canada



(b) Mexico

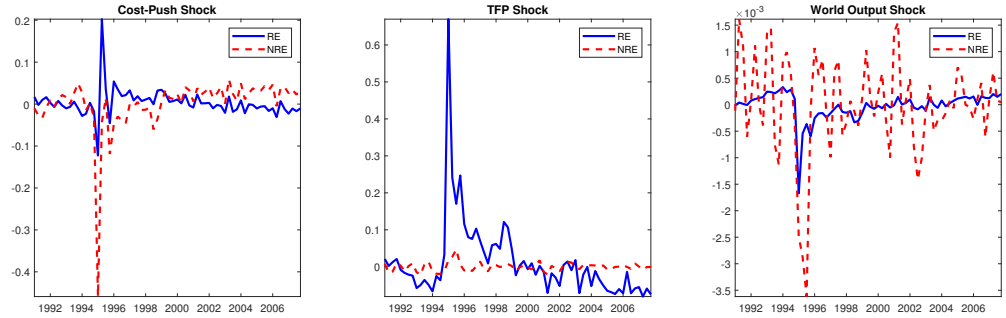
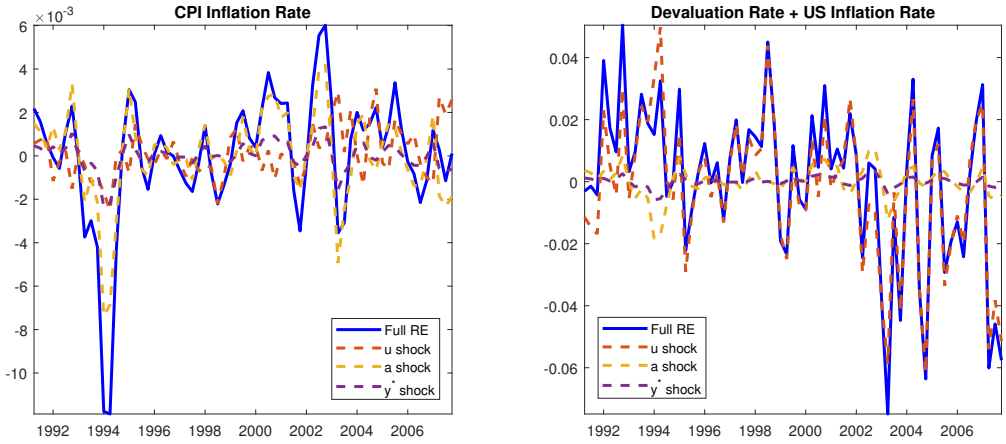


Figure C.25: Historical Decomposition of Observables, Canada

(a) RE Model



(b) NRE Model

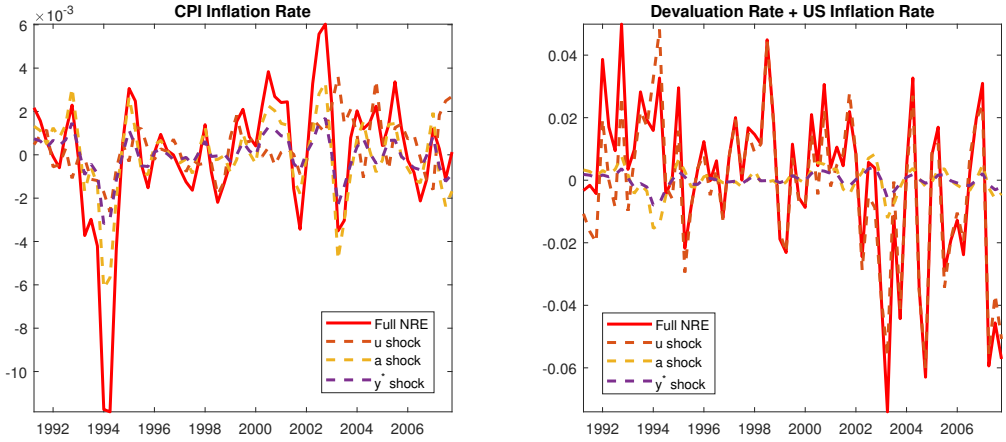
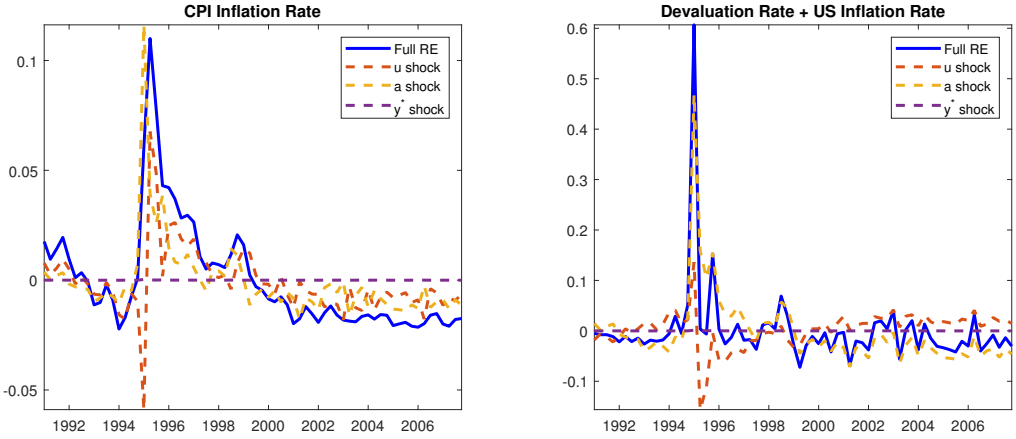


Figure C.26: Historical Decomposition of Observables, Mexico

(a) RE Model



(b) NRE Model

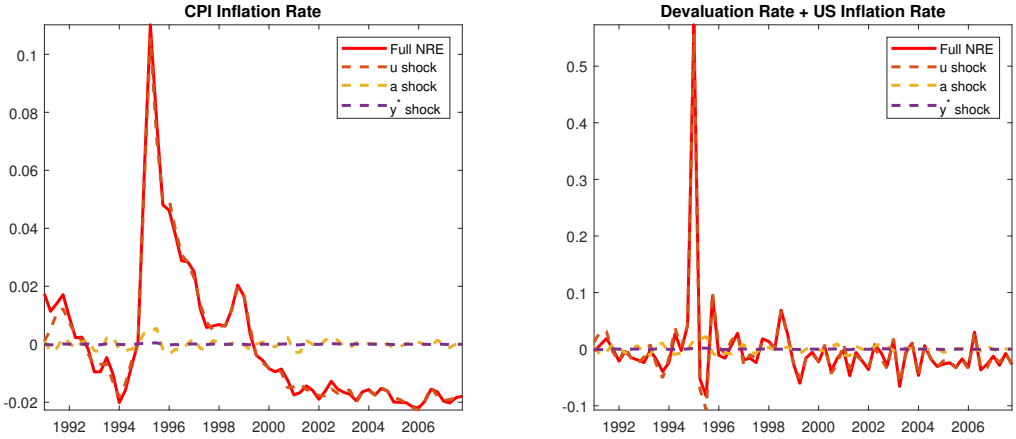


Figure C.27: Decomposition of Model-Predicted Domestic Inflation, Devaluation, and CPI Inflation Driven by Optimal Policy Intervention to Cost-Push Shock, Canada

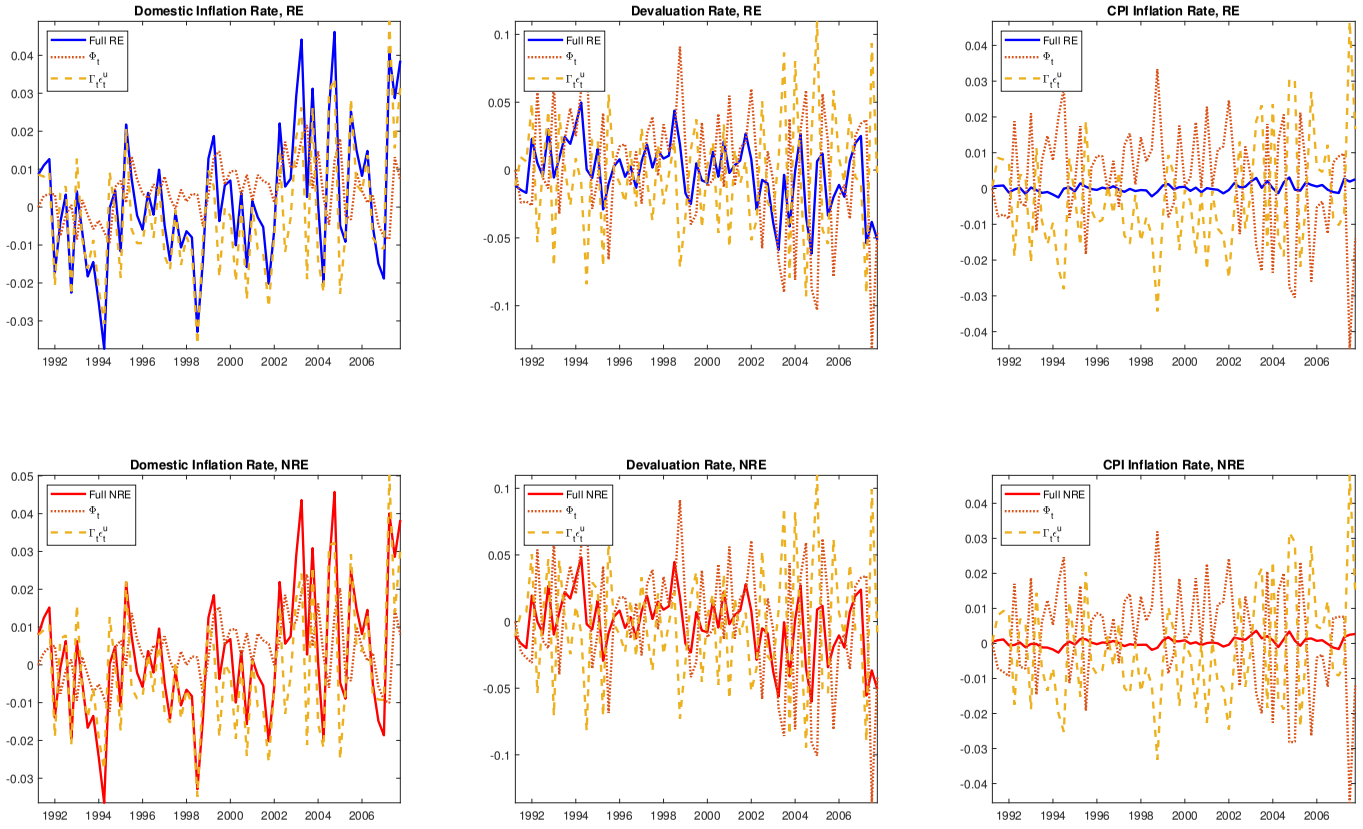


Figure C.28: Decomposition of Model-Predicted Domestic Inflation, Devaluation, and CPI Inflation Driven by Optimal Policy Intervention to Cost-Push Shock, Mexico