Policy Rule Regressions with Survey Data†

Leifei Lyu, Fei Tan, and Zheliang Zhu*

March 14, 2025

Abstract

Endogeneity bias is ubiquitous and can be sizable in the ordinary least squares (OLS) estimation of policy rule coefficients. This paper introduces a simple procedure that leverages survey data to rectify the bias under flexible information assumptions. We decompose policy rule regressors (e.g., inflation and output gap) into their forecasts made before policy decisions and the associated forecast errors. The forecasts are readily available in survey data and, by construction, orthogonal to the forecast errors and policy shocks under complete information. We further orthogonalize the forecast errors to remove the bias in the presence of information rigidities. Using Monte Carlo simulations, we showcase the efficacy of this procedure in a prototypical new Keynesian model under distinct information settings. As an empirical application, we employ Bayesian methods to compare the performance of the standard OLS approach using real-time data and our approach based on survey data in estimating monetary and fiscal policy rules. Marginal likelihood estimates reveal that our approach consistently outperforms across nearly all model specifications considered.

Keywords: Endogeneity bias; Incomplete information; Two-stage OLS.

JEL Classification: C11, C22, E52

---

†We thank Francisco Buera, SangMok Lee, Rodolfo Manuelli, and Ping Wang as well as seminar participants at Washington University in St. Louis, Insightful Minds in International Macro, 2023 EAC ISBA Conference, 2023 Midwest Macroeconomics Group Meetings, and 33rd Annual Meeting of the Midwest Econometrics Group for helpful comments.

*Lyu: Department of Economics, Washington University in St. Louis, email: leifei.lyu@wustl.edu; Tan: Department of Economics, Chaifetz School of Business, Saint Louis University and Center for Economic Behavior and Decision-Making, Zhejiang University of Finance and Economics, email: tanf@slu.edu; Zhu: Institute of Chinese Financial Studies, Southwestern University of Finance and Economics, email: zhuzheliang@swufe.edu.cn.

1 Introduction

The policy reaction functions of governments and central banks are often summarized by straightforward rules. A prominent illustration is the Taylor rule, delineated in Taylor (1993), which posits that the central bank sets short-term interest rates contingent upon variations in the inflation rate and the output gap. However, estimating such reaction functions using ordinary least squares (OLS) often introduces an endogeneity bias. This bias occurs because central banks adjust their policies in response to economic variables that are simultaneously influenced by policy shocks.

Under a complete-information framework, Carvalho, Nechio and Tristao (2021) argue in favor of OLS estimation of policy rules despite this endogeneity bias. Nevertheless, a growing body of work examines the implications of departures from the full-information setting. Coibion and Gorodnichenko (2015) and Angeletos, Huo and Sastry (2021), for instance, provide empirical evidence supporting the presence of incomplete information. These information frictions can introduce additional biases into the estimated coefficients, raising the natural question of how to mitigate such biases under flexible information settings.

We introduce a simple method using survey data that effectively reduces the endogeneity bias. Under complete information, we decompose policy rule regressors (e.g., inflation and output gap) into two components: forecasts derived from the previous period and their associated forecast errors. The forecasts, by construction, are orthogonal to both the forecast errors and policy shocks under complete information. This orthogonality arises because forecast errors reflect current period information, whereas forecasts are based on previous period information. We refer to this approach as the one-step method.

Although survey data forecasters are among the most informed economic agents, information rigidities may exist. If agents form expectations under incomplete information, the average forecast errors across agents are not orthogonal to the average forecasts of policy rule regressors, which reintroduces endogeneity bias when employing our one-step method. In this case, we further orthogonalize the forecast errors using the empirical specification suggested by Coibion and Gorodnichenko (2015) (CG henceforth). We refer to this approach as the two-step method. In the first step, we employ a regression of forecast errors on ex-ante mean forecast revisions to obtain their predicted values. In the second step, we create new regressors by summing these predicted values and the forecasts, ensuring the new regressors are orthogonal to the error term.

Our approach offers several advantages. First, it addresses the endogeneity bias in estimating policy rules while yielding robust estimates regardless of whether policymakers rely on contemporaneous data or forecasts in decision-making [see also Orphanides (2003) and Coibion and Gorodnichenko (2012)]. Second, forecasts for key macroeconomic variables are readily available from the U.S. Survey of Professional Forecasters (SPF) and the Tealbook Data Set. Third, our method leverages earlier period information, thereby avoiding the common approach of using lagged variables as instrumental variables. The latter approach may pose challenges when policy rules exhibit inertia because the lagged variables may be included multiple times within the rules, potentially leading to a multicollinearity issue.

To assess the efficacy of our method, we first quantify the estimation bias in dynamic stochastic general equilibrium (DSGE) models and investigate how survey data can help reduce this bias under various information settings. Using a standard new Keynesian model under complete information, our simulation results reveal that the conventional OLS approach can yield highly inaccurate estimates, while the use of survey data substantially improves these estimates. We also examine a sticky information model based on Reis (2009), where only a fraction of agents update their information each period. The simulation results demonstrate that the OLS approach using real-time data leads to biased estimates, whereas our methods significantly improve the estimates, with the two-step method outperforming the one-step approach.

We then conduct empirical analysis using a Bayesian framework to compare the performance of standard OLS with our proposed methods for estimating monetary and fiscal policy rules. Using inflation and output gap forecasts from the SPF and Tealbook dataset, our sample covers the period from the fourth quarter of 1987 to the fourth quarter of 2006, just before the Great Recession in 2007.[1] During this period, we cannot reject the null hypothesis of complete information, implying that information rigidities do not pose a serious issue. Therefore, the one-step method is sufficient for our purposes. The Bayesian results show that our approach consistently yields higher marginal likelihood compared to standard OLS estimation across nearly all model specifications, with the one-step method outperforming the two-step method in most specifications.

The rest of the paper is organized as follows. In Section 2, we introduce our method to address the endogeneity issue in policy rule regressions. Section 3 quantifies how the proposed method reduces the endogeneity bias under different information settings. Section 4 presents the empirical applications of our method. Section 5 concludes the paper.

2 Policy Rule Regressions

In this section, we propose empirical methods that use survey data in estimating policy rules to eliminate the endogeneity bias under various information settings. Consider the following general autoregressive processes

$$y_t = \sum_{k=1}^{K} \rho_k y_{t-k} + \sum_{i=1}^{I} \beta_i X_{i,t} + \varepsilon_t, \tag{2.1}$$

where y_t represents the policy choice, ρ_k measures the policy smoothing on the k-th lagged term, β_i represents the coefficient on the controlled variable $X_{i,t}$, and the exogenous shock ε_t is assumed

[1] The Great Recession is widely considered as a structural break in the literature [see, e.g., Bunzel and Enders (2010)].

to be independently and identically distributed. An endogeneity issue may arise when estimating β_i using OLS regression, as the policy rule regressor $X_{i,t}$ could be correlated with the error term ε_t. To address this endogeneity issue, we decompose the policy rule regressors to separate the new controlled variables from the new error term based on the timing of information arrival. This decomposition ensures that the two components remain orthogonal.

Specifically, we decompose $X_{i,t}$ into two components, i.e., forecasts and the associated forecast errors, and rewrite (2.1) as

$$y_t = \sum_{k=1}^{K} \rho_k y_{t-k} + \sum_{i=1}^{I} \beta_i \mathbb{E}_{t-1} X_{i,t} + \underbrace{\sum_{i=1}^{I} \beta_i [X_{i,t} - \mathbb{E}_{t-1} X_{i,t}] + \varepsilon_t}_{\text{new error term}}, \tag{2.2}$$

where $\mathbb{E}_{t-1}X_{i,t}$ represents the ex-ante forecast of $X_{i,t}$ and $X_{i,t} - \mathbb{E}_{t-1}X_{i,t}$ captures the corresponding ex-post forecast error. It is important to note that ex-ante forecasts are formed using information available at period $t-1$ and earlier. In other words, $\mathbb{E}_{t-1}X_{i,t}$ is independent of any information from period t. Meanwhile, the forecast error is linked to information available at time t. Similarly, ε_t also depends on time t information. Therefore, the new error term, $\sum_{i=1}^{I} \beta_i[X_{i,t} - \mathbb{E}_{t-1}X_{i,t}] + \varepsilon_t$, is determined exclusively by information available at time t. This decomposition effectively helps mitigate the endogeneity issue in parameter estimation.

We consider the Taylor rule as an example. In a Taylor-type interest rate reaction function, monetary policymakers respond to real-time inflation and output gap as follows

$$r_t = \psi_1 \pi_t + \psi_2 ygap_t + \varepsilon_t, \tag{2.3}$$

where r_t denotes the nominal interest rate, π_t represents the inflation rate, $ygap_t$ is the output gap at time t, and ψ_1 and ψ_2 are the policy coefficients that need to be estimated. Given that both π_t and $ygap_t$ may be correlated with ε_t, we apply the decomposition in (2.2) and rewrite (2.3) as

$$r_t = \psi_1 \mathbb{E}_{t-1}\pi_t + \psi_2 \mathbb{E}_{t-1} ygap_t + \underbrace{\psi_1[\pi_t - \mathbb{E}_{t-1}\pi_t] + \psi_2[ygap_t - \mathbb{E}_{t-1}ygap_t] + \varepsilon_t}_{\text{new error term}}, \tag{2.4}$$

where $\mathbb{E}_{t-1}\pi_t$ and $\mathbb{E}_{t-1}ygap_t$ represent the inflation and output gap ex-ante forecasts based on information available at $t-1$, respectively.

Given this specification, we can use survey data for parameter estimation, as it provides direct measures of forecasts.[2] For example, inflation forecasts are available from the Survey of Professional Forecasters, a survey of approximately 40 experts conducted by the Federal Reserve Bank

[2] Our specification (2.4) is fundamentally different from Coibion and Gorodnichenko (2012), which assumes that policymakers target forecasts instead of real-time variables.

of Philadelphia. Moreover, output gap forecasts can be obtained from the Tealbook (formerly the Greenbook) dataset, which contains real-time estimates and projections of the output gap used by the Board of Governors of the Federal Reserve System. We refer to this approach as the one-step method.

However, the orthogonality between the regressors and the error term in (2.2) may break down in the presence of information rigidity. As CG demonstrate, ex-post forecast errors can be correlated with ex-ante forecast revisions.[3] In essence, average forecast errors across agents may reflect information from the previous period, creating a correlation between the error term and ex-ante forecasts. As a result, applying the OLS method to (2.2) may fail to eliminate the endogeneity bias.

To address this issue, we further orthogonalize average forecast errors on average forecast revisions. Specifically, we estimate

$$\underbrace{X_{i,t} - \mathbb{E}_{t-1}X_{i,t}}_{\Delta X_{i,t}} = \alpha_i + \beta_i(\mathbb{E}_{t-1}X_{i,t} - \mathbb{E}_{t-2}X_{i,t}) + \varepsilon_{i,t}. \tag{2.5}$$

We then obtain the predicted value of $\Delta X_{i,t}$, denoted as $\Delta\widehat{X}_{i,t}$, and rewrite (2.2) as

$$y_t = \sum_{k=1}^{K} \rho_k y_{t-k} + \sum_{i=1}^{I} \beta_i \underbrace{\left[\mathbb{E}_{t-1}X_{i,t} + \Delta\widehat{X}_{i,t}\right]}_{\text{new policy rule regressor}} + \varepsilon_t, \tag{2.6}$$

where the new policy rule regressor, $\mathbb{E}_{t-1}X_{i,t} + \Delta\widehat{X}_{i,t}$, is by construction based on information available at time $t-1$, while the error term ε_t remains independent of information at $t-1$. The endogeneity bias is thus removed, allowing us to use the OLS method to estimate the coefficients in (2.6). We refer to this approach as the two-step method. In the first step, we regress ex-post forecast errors on ex-ante forecast revisions to obtain their predicted values. In the second step, we construct new policy rule regressors by summing these predicted forecast errors with the ex-ante forecasts. Finally, we estimate the policy rule using these newly constructed regressors.

3 Simulations

In this section, we quantify the endogeneity bias in single-equation estimations of policy rules within DSGE models under various information settings. Our simulation results show that the one-step method significantly reduces the endogeneity bias under complete information, while the two-step method further mitigates the bias in the presence of information rigidity.

[3] CG propose a novel approach to test the complete information hypothesis. They introduce two models of information rigidity: the sticky-information model and the noisy-information model. Both models exhibit a similar empirical relationship between ex-post mean forecast errors and ex-ante mean forecast revisions.

3.1 Complete-Information Model We first demonstrate how using survey data can reduce the endogeneity bias in a complete-information setting. We consider a textbook version of the new Keynesian model presented in Woodford (2003a) and Galí (2008), augmented with a fiscal policy rule. The model consists of a representative household and a continuum of firms, each producing a differentiated product; only a fraction of firms can adjust their prices each period; a cashless economy with one-period nominal bonds B_t selling at a price of $1/R_t$, where R_t is the monetary policy instrument; a primary surplus s_t with lump-sum taxation and zero government spending, resulting in consumption being equal to output, $c_t = y_t$; a monetary authority and fiscal authority.

Let $\hat{x}_t \equiv \ln(x_t) - \ln(x)$ represent the log-deviation of a generic variable x_t from its steady state x. A log-linear approximation of the model's equilibrium conditions around the zero-inflation steady state yields

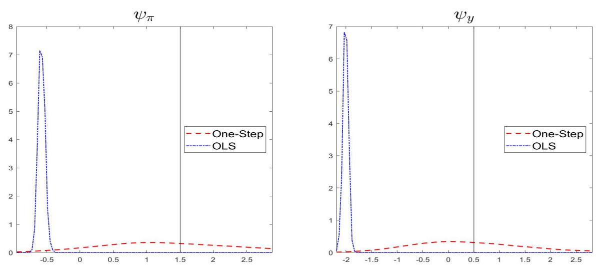
$$\begin{aligned}
\text{Dynamic IS equation:} \quad & \hat{y}_t = \mathbb{E}_t\hat{y}_{t+1} - \tau^{-1}(\hat{R}_t - \mathbb{E}_t\hat{\pi}_{t+1} - \mathbb{E}_t\hat{z}_{t+1}) && (3.1)\\
\text{New Keynesian Phillips curve:} \quad & \hat{\pi}_t = \beta\mathbb{E}_t\hat{\pi}_{t+1} + \kappa\hat{y}_t + \kappa\tau^{-1}\hat{u}_t && (3.2)\\
\text{Monetary policy:} \quad & \hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)(\psi_\pi\hat{\pi}_t + \psi_y\hat{y}_t) + \epsilon_{R,t} && (3.3)\\
\text{Fiscal policy:} \quad & \hat{s}_t = \rho_s\hat{s}_{t-1} + (1-\rho_s)\psi_s\hat{b}_{t-1} + \epsilon_{s,t} && (3.4)\\
\text{Government budget constraint:} \quad & \hat{b}_t = \beta^{-1}(\hat{b}_{t-1} + \hat{R}_{t-1} - \hat{\pi}_t - \hat{z}_t) - (\beta^{-1}-1)\hat{s}_t && (3.5)\\
\text{Technology shock:} \quad & \hat{z}_t = \rho_z\hat{z}_{t-1} + \epsilon_{z,t} && (3.6)\\
\text{Markup shock:} \quad & \hat{u}_t = \rho_u\hat{u}_{t-1} + \epsilon_{u,t} && (3.7)
\end{aligned}$$

where $\pi_t = P_t/P_{t-1}$ is the inflation between periods $t-1$ and t, and $b_t = B_t/P_t$ is the real debt at the end of period t. Equations (3.1)–(3.3) form the key building blocks of the standard new Keynesian model, (3.4) is the model analog to many surplus-debt regression studies that aim to test for fiscal sustainability, and (3.5) is the log-linearized version of the government's flow budget identity, $\frac{1}{R_t}\frac{B_t}{P_t} + s_t = \frac{B_{t-1}}{P_t}$. Together, (3.1)–(3.5) form a linear expectational difference equation system, capturing the core dynamics of most monetary DSGE models.

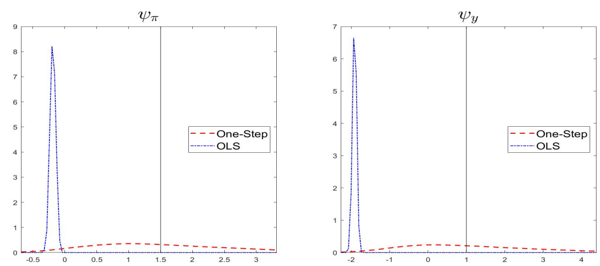
To evaluate the effectiveness of our one-step method relative to traditional OLS in estimating policy parameters, we examine the results based on simulated data under different regimes of the monetary policy parameters ψ_π and ψ_y, while keeping other structural parameters at standard values from the literature (see Table 1). Figure 1 reports the simulation results across regimes. The dashed-dotted blue line depicts the distribution of parameter estimates obtained using OLS with real-time variables, $\hat{y}_t$ and $\hat{\pi}_t$. The dashed red line represents the distribution derived from our one-step method using ex-ante forecasts, $\mathbb{E}_{t-1}\hat{y}_t$ and $\mathbb{E}_{t-1}\hat{\pi}_t$. Each simulation was repeated 500 times and the vertical black solid lines indicate the true parameter values.

Table 1: Parameters

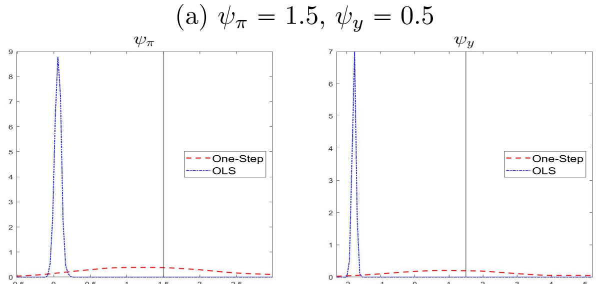
Parameter	Description	value
τ	elasticity of intertemporal substitution	2.00
κ	the slope of the new Keynesian Philips curve	0.20
β	discount factor	0.99
ψ_s	coefficients of fiscal policy on output gap	0.5
ρ_R	monetary policy smoothing	0.5
ρ_s	fiscal policy smoothing	0.9
ρ_z	AR(1) coefficient on technology shock	0.6
ρ_u	AR(1) coefficient on Markup shock	0.7
σ_R	standard deviations on monetary policy shock	1.2
σ_s	standard deviations on fiscal policy shock	1.8
σ_z	standard deviations on technology shock	0.4
σ_u	standard deviations on markup shock	1.5



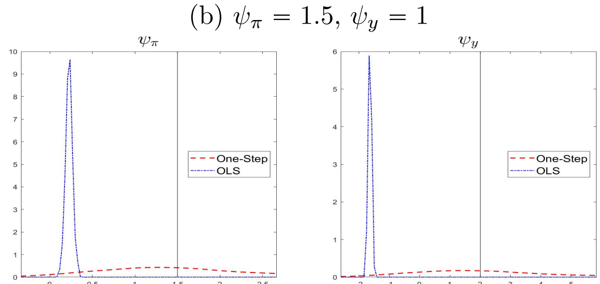
(a) $\psi_\pi = 1.5$, $\psi_y = 0.5$



(b) $\psi_\pi = 1.5$, $\psi_y = 1$



(c) $\psi_\pi = 1.5$, $\psi_y = 1.5$



(d) $\psi_\pi = 1.5$, $\psi_y = 2$

Figure 1: Distribution of 500 simulations with full information

Notes: This figure illustrates the distribution of estimated policy coefficients (ψ_π and ψ_y) obtained using different methods across various regimes. The dashed-dotted blue line represents the distribution from the OLS method, while the dashed red line corresponds to the one-step method.

The results demonstrate that the mean of the one-step estimator is consistently close to the true value. Therefore, our simulations confirm the effectiveness of the decomposition in (2.4), indicating that estimation using survey data outperforms real-time data across all regimes and enhances the accuracy of OLS estimates.[4] Furthermore, our results reveal that OLS estimates can deviate substantially from the true values, suggesting a need for further examination of the conclusion in Carvalho, Nechio and Tristao (2021), who advocate for the use of OLS in Taylor rule estimation despite the presence of endogeneity bias.

3.2 Sticky-Information Model

Due to information rigidity, ex-post forecast errors may be correlated with ex-ante forecasts, leading to the potential endogeneity bias in the one-step method using survey data. To address this issue, we adopt the two-step method, as outlined in (2.5), to construct new policy rule regressors and further mitigate the bias.

We consider the sticky-information model as presented in Reis (2009).[5] In the goods markets, inattentive firms update their information at rate λ, while $1-\lambda$ fraction of firms keep one-period old information. The core distinction between the complete information and the sticky information models stems from the following Phillips curve

$$p_t = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i \mathbb{E}_{t-i} \left\{ \frac{\beta(w_t - p_t) + (1-\beta)y_t - a_t}{\beta + \nu(1-\beta)} - \frac{\beta}{(\nu-1)[\beta+\nu(1-\beta)]}\nu_t \right\}, \tag{3.8}$$

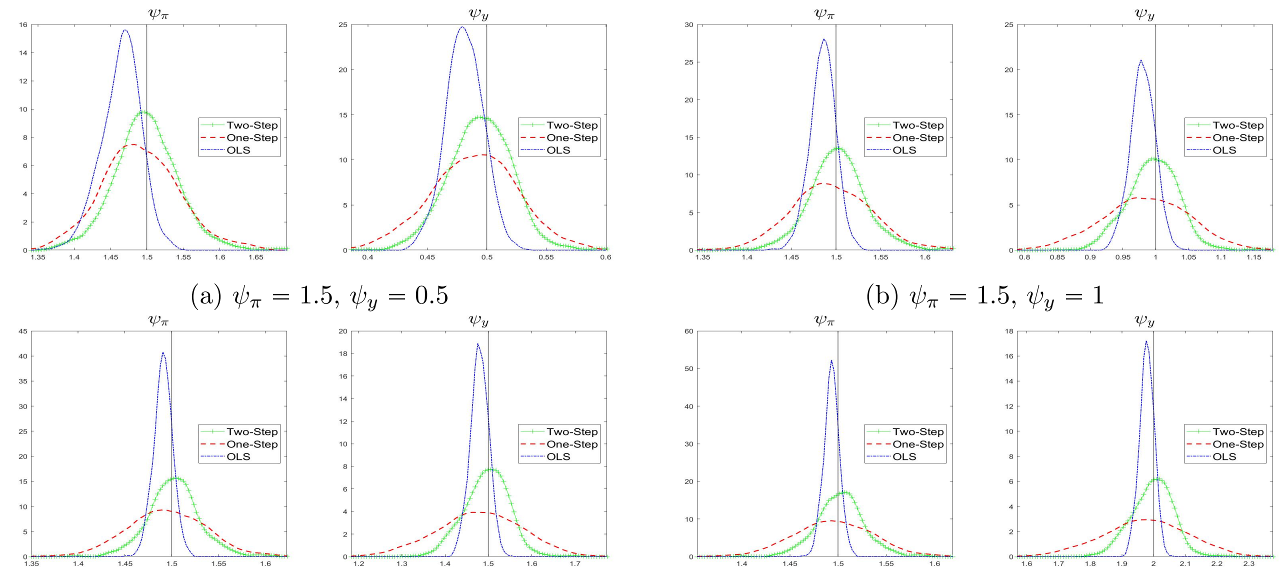
where p_t is the price level, $w_t - p_t$ is the real wage, y_t is the output, a_t represents the productivity shock, and ν_t denotes the random substitutability of varieties of goods with ν being the steady-state elasticity of substitution for goods. The parameter β denotes the labor share. Other structural equations in Reis (2009) include an IS curve, a wage-setting equation, a conventional production function, and a Taylor rule for monetary policy.

We adopt the solution method proposed by Verona and Wolters (2014) and conduct 500 simulations using their calibrated parameters. Figure 2 compares the performance of OLS estimation, the one-step method, and the two-step method. The dashed-dotted blue line represents the distribution from the OLS method, the dashed red line corresponds to the one-step method, and the green line with plus signs illustrates the distribution from the two-step method. The results show that the performance of the one-step method deteriorates significantly in the presence of information rigidity. This occurs because ex-post forecast errors are correlated with ex-ante forecasts, undermining the effectiveness of directly using survey data for parameter estimation. In contrast, the two-step method, which projects average forecast errors onto average forecast revisions as outlined in (2.5), effectively mitigates the endogeneity bias introduced by sticky information. As a result, the mean of the two-step estimator distribution is consistently closer to

[4] The variance of the estimates is relatively large, which aligns with intuition, as survey data inherently exhibit greater uncertainty compared to real-time data. Nevertheless, the findings show that, despite significant heterogeneity in individual information reception, estimation using survey data remains highly effective on average.

[5] We use sticky-information to characterize information rigidity because the endogeneity problem in parameter estimation stems from the interaction of information across different time periods.

the true value across all regimes.[6]



(a) $\psi_\pi = 1.5$, $\psi_y = 0.5$

(b) $\psi_\pi = 1.5$, $\psi_y = 1$

(c) $\psi_\pi = 1.5$, $\psi_y = 1.5$

(d) $\psi_\pi = 1.5$, $\psi_y = 2$

Figure 2: Distribution of 500 simulations in a sticky information model

Notes: This figure illustrates the distribution of estimated policy coefficients (ψ_π and ψ_y) obtained through various methods. The dashed-dotted blue line represents the distribution from the OLS method, the dashed red line corresponds to the one-step method, and the green line marked with plus signs illustrates the distribution from the two-step method.

4 Empirical Analysis

In this section, we apply our proposed methods to estimate monetary and fiscal policy rules and compare their performance with traditional OLS estimates. To rigorously evaluate our methods, we employ the Bayesian model scan approach developed by Chib and Zeng (2020), which identifies the optimal model by comparing the log marginal likelihoods of candidate empirical models.

Specifically, the parameters of the general empirical specification (2.1) are given as

$$\boldsymbol{\theta} = (\rho_1, \ldots, \rho_K, \beta_1, \ldots, \beta_I, \sigma). \tag{4.1}$$

Assume that we have J models denoted by $\mathcal{M}_1, \ldots, \mathcal{M}_J$. Let $\Pr(\mathcal{M}_j)$ represent the prior probability of model $\mathcal{M}_j$ on the model space and $m(\mathbf{y}_{1:T}|\mathcal{M}_j, \mathbf{X}_{1:T})$ denote the marginal likelihood of model $\mathcal{M}_j$, where $\mathbf{y}_{1:T}$ is the sample data for the dependent variable and $\mathbf{X}_{1:T}$ corresponds to the sample data for the controlled variables over T time periods. According to the Bayes theorem, the posterior probability of $\mathcal{M}_j$ is given by

[6]Consistent with the complete-information case, the variance of estimates using survey data remains higher than that using real-time data, reflecting the greater uncertainty inherent in expectational variables.

$$\Pr(\mathcal{M}_j|y_{1:T}, X_{1:T}) = \frac{\Pr(\mathcal{M}_j)m(\mathbf{y}_{1:T}|\mathcal{M}_j, \mathbf{X}_{1:T})}{\sum_{l=1}^{J}\Pr(\mathcal{M}_l)m(\mathbf{y}_{1:T}|\mathcal{M}_l, \mathbf{X}_{1:T})}, \tag{4.2}$$

where $m(\mathbf{y}_{1:T}|\mathcal{M}_j, \mathbf{X}_{1:T})$ is calculated numerically using the method proposed by Chib (1995). In Bayesian analysis, models with higher marginal likelihoods are preferred.[7]

We consider the following two types of applications, i.e., Taylor-type interest rate rules and government spending rules, given by

$$r_t = \sum_{i=1}^{I}\rho_i r_{t-i} + \psi_1\pi_t + \psi_2 ygap_t + \varepsilon_t, \tag{4.3}$$

and

$$g_t = \sum_{i=1}^{I}\varrho_i g_{t-i} + \phi_1 b_{t-1} + \phi_2 ygap_t + \varepsilon_t, \tag{4.4}$$

where g_t denotes the government spending and b_t represents the government debt. To evaluate whether policy rules exhibit inertia, we allow I to range from zero to two. We consider three model specifications, each of which can be estimated using OLS and our proposed methods. This results in a total of nine models for comparison. The first set of models corresponds to the case of $I = 0$, meaning policy rules exclude lagged terms. These models, estimated using OLS, the one-step method, and the two-step method, are denoted as $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$, respectively. The second set of models allows for $I = 1$, incorporating a one-period lag in policy rules, with the corresponding estimates labeled $\mathcal{M}_4$, $\mathcal{M}_5$, and $\mathcal{M}_6$. Finally, the third set of models considers $I = 2$, where policy rules include a two-period lag, with the relevant estimates labeled $\mathcal{M}_7$, $\mathcal{M}_8$, and $\mathcal{M}_9$.

4.1 MCMC Simulations

We use the initial 15% of the total sample, spanning from 1987:IV to 2006:IV, to form the prior. We sample 10,000 draws from the posterior distribution, discarding the first 1,000 as burn-in draws. To test the complete-information hypothesis, we estimate specification (2.5) for inflation and output gap expectations, as shown in Table 2. The table reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th quantiles of the estimated parameters. A coefficient is considered statistically significant if the 2.5th and 97.5th percentiles share the same sign. Over the sample period from the fourth quarter of 1987 to the fourth quarter of 2006, we cannot reject the complete-information hypothesis, as the 2.5th and 97.5th percentiles exhibit opposite signs for all estimates. Based on the theoretical framework in Section 2, we expect the one-step method to outperform the two-step method.

[7] Suppose that the prior probabilities are equal. Let δ_{lm} denote the difference in the log marginal likelihoods between models $\mathcal{M}_l$ and $\mathcal{M}_m$, which is calculated as $\ln m(y_{1:T}|\mathcal{M}_l, X_{1:T}) - \ln m(y_{1:T}|\mathcal{M}_m, X_{1:T})$. In a two-model case, the posterior probability of model l is greater than 67%, 75%, and 80%, which requires that δ_{lm} is greater than 0.694676, 1.098612, and 1.386294, respectively.

Table 2: Tests of inflation and output gap expectations processes

Inflation				Output gap			
Name	postmean	postsd	[0.025 0.975]	Name	postmean	postsd	[0.025 0.975]
α_π	-0.048	0.082	[-0.209 0.112]	α_y	0.026	0.066	[-0.104 0.154]
β_π	0.249	0.207	[-0.157 0.664]	β_y	-0.099	0.097	[-0.290 0.095]
σ^2	0.580	0.100	[0.416 0.805]	σ^2	0.406	0.071	[0.290 0.564]

Note: The table reports coefficient estimates for the specification (2.5) with respect to expectations of inflation and output gap. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters.

Tables 3–8 present the estimates of monetary and fiscal policy rules specified in (4.3) and (4.4). These estimates include the posterior mean, posterior standard deviation, and 2.5th and 97.5th quantiles of the parameters. To facilitate comparison across policy rule specifications, we rescale certain parameters, denoted with an asterisk. For instance, in $\mathcal{M}_6$, we compute $\psi_1^* = \psi_1/(1-\rho_1-\rho_2)$. A detailed discussion of these policy rule estimates will follow in subsequent sections.

4.2 Monetary Policy Rule Estimations Tables 3–5 compare OLS estimations using real-time data with the one-step and two-step methods using survey data across various specifications of monetary policy rules. Our findings reveal that both the one-step and two-step methods yield higher marginal likelihoods than OLS estimations under all specifications. Since the null hypothesis of complete information from 1987 to 2006 cannot be rejected, the one-step method consistently outperforms the two-step method across all specifications. Additionally, our results indicate that the OLS method underestimates the coefficients on inflation, while the estimates for output gap remain inconclusive.

Previous studies using quarterly data to estimate monetary policy rules have highlighted the presence of policy inertia. In our empirical analysis, the coefficient on the one-period lagged federal funds rate is substantial (0.870, 0.724, and 0.758), with the 2.5th and 97.5th percentiles significantly different from zero. This suggests that interest rate smoothing significantly improves the log marginal likelihoods for the OLS estimation, the one-step method, and the two-step method, raising the values from -143.90, -94.90, and -105.80 to -33.53, -23.54, and -25.64, respectively. Moreover, the coefficients on the second-order interest rate in $\mathcal{M}_7$, $\mathcal{M}_8$, and $\mathcal{M}_9$ are negative and significantly below zero. Incorporating the second-order interest rate further increases the log marginal likelihoods to -18.47, -14.64, and -16.96, respectively. These findings align with the conclusions in Coibion and Gorodnichenko (2012) and Woodford (2003b), suggesting that the central bank engaged in interest rate smoothing during the sample period.

Table 3: Estimates of monetary policy without lag

$\mathcal{M}_1$	postmean	postsd	[0.025	0.975]
$\psi_1 : \pi_t$	1.893	0.103	[1.692	2.092]
$\psi_2 : y_t$	0.661	0.130	[0.406	0.922]
σ^2	6.809	1.059	[5.042	9.141]
Log marginal likelihood: -143.9046				
$\mathcal{M}_2$	postmean	postsd	[0.025	0.975]
$\psi_1 : \mathbb{E}_{t-1}\pi_t$	1.998	0.048	[1.903	2.092]
$\psi_2 : \mathbb{E}_{t-1}y_t$	0.702	0.060	[0.584	0.821]
σ^2	1.261	0.215	[0.908	1.740]
Log marginal likelihood: -94.90397				
$\mathcal{M}_3$	postmean	postsd	[0.025	0.975]
$\psi_1 : \mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1}$	2.026	0.059	[1.910	2.140]
$\psi_2 : \mathbb{E}_{t-1}y_t + \widehat{\Delta y}_{t,t-1}$	0.691	0.072	[0.550	0.836]
σ^2	1.832	0.308	[1.325	2.518]
Log marginal likelihood: -105.7976				

Note: The table reports coefficient estimates for the specification (4.3) with $I = 0$. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters. The results indicate that all estimates are significant throughout the sample period. Models $\mathcal{M}_1$ - $\mathcal{M}_3$ correspond to estimates obtained using the OLS with real-time data, the one-step method with survey data, and the two-step method with survey data, respectively. Among these estimates, the one-step method yields the highest marginal likelihood.

Table 4: Estimates of monetary policy with one-period lag

$\mathcal{M}_4$	postmean	postsd	[0.025	0.975]
$\rho_1 : R_{t-1}$	0.870	0.024	[0.823	0.917]
$\psi_1 : \pi_t$	0.258	0.049	[0.162	0.353]
$\psi_2 : y_t$	0.167	0.027	[0.113	0.221]
$\psi_1^* : (1-\rho_1)\pi_t$	1.983	0.156	[1.685	2.303]
$\psi_2^* : (1-\rho_1)y_t$	1.310	0.245	[0.908	1.873]
σ^2	0.174	0.030	[0.124	0.241]
Log marginal likelihood: -33.52841				
$\mathcal{M}_5$	postmean	postsd	[0.025	0.975]
$\rho_1 : R_{t-1}$	0.724	0.033	[0.659	0.789]
$\psi_1 : \mathbb{E}_{t-1}\pi_t$	0.554	0.068	[0.421	0.687]
$\psi_2 : \mathbb{E}_{t-1}y_t$	0.252	0.029	[0.195	0.309]
$\psi_1^* : (1-\rho_1)\mathbb{E}_{t-1}y_t$	2.005	0.060	[1.886	2.125]
$\psi_2^* : (1-\rho_1)\mathbb{E}_{t-1}y_t$	0.916	0.083	[0.765	1.091]
σ^2	0.126	0.022	[0.090	0.175]
Log marginal likelihood: -23.53923				
$\mathcal{M}_6$	postmean	postsd	[0.025	0.975]
$\rho_1 : R_{t-1}$	0.758	0.031	[0.698	0.819]
$\psi_1 : \mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1}$	0.494	0.064	[0.366	0.620]
$\psi_2 : \mathbb{E}_{t-1}y_t + \widehat{\Delta y}_{t,t-1}$	0.228	0.028	[0.172	0.283]
$\psi_1^* : (1-\rho_1)(\mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1})$	2.041	0.073	[1.898	2.187]
$\psi_2^* : (1-\rho_1)(\mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1})$	0.946	0.099	[0.768	1.158]
σ^2	0.135	0.024	[0.097	0.188]
Log marginal likelihood: -25.64392				

Note: The table reports coefficient estimates for the specification (4.3) with $I = 1$. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters. The results indicate that all estimates are significant throughout the sample period. Models $\mathcal{M}_4$ - $\mathcal{M}_6$ correspond to estimates obtained using the OLS with real-time data, the one-step method with survey data, and the two-step method with survey data, respectively. Among these estimates, the one-step method yields the highest marginal likelihood. The rescaled parameters are defined as $\psi_1^* \equiv \psi_1/(1-\rho_1)$ and $\psi_2^* \equiv \psi_2/(1-\rho_1)$.

Table 5: Estimates of monetary policy with two-period lag

$\mathcal{M}_7$	postmean	postsd	[0.025 0.975]
$\rho_1 : R_{t-1}$	1.465	0.095	[1.273 1.651]
$\rho_2 : R_{t-2}$	−0.550	0.086	[−0.716 − 0.380]
$\psi_1 : \pi_t$	0.161	0.041	[0.081 0.242]
$\psi_2 : y_t$	0.086	0.025	[0.038 0.135]
$\psi_1^* : (1-\rho_1-\rho_2)\pi_t$	1.908	0.203	[1.515 2.316]
$\psi_2^* : (1-\rho_1-\rho_2)y_t$	1.032	0.282	[0.560 1.638]
σ^2	0.107	0.019	[0.076 0.150]
Log marginal likelihood: -18.4747			
$\mathcal{M}_8$	postmean	postsd	[0.025 0.975]
$\rho_1 : R_{t-1}$	1.249	0.114	[1.021 1.467]
$\rho_2 : R_{t-2}$	−0.437	0.091	[−0.614 − 0.255]
$\psi_1 : \mathbb{E}_{t-1}\pi_t$	0.370	0.071	[0.235 0.509]
$\psi_2 : \mathbb{E}_{t-1}y_t$	0.150	0.033	[0.086 0.215]
$\psi_1^* : (1-\rho_1-\rho_2)\mathbb{E}_{t-1}\pi_t$	1.970	0.082	[1.800 2.131]
$\psi_2^* : (1-\rho_1-\rho_2)\mathbb{E}_{t-1}y_t$	0.799	0.107	[0.597 1.013]
σ^2	0.095	0.017	[0.067 0.133]
Log marginal likelihood: -14.64363			
$\mathcal{M}_9$	postmean	postsd	[0.025 0.975]
$\rho_1 : R_{t-1}$	1.297	0.118	[1.061 1.522]
$\rho_2 : R_{t-2}$	−0.453	0.096	[−0.639 − 0.260]
$\psi_1 : \mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1}$	0.312	0.068	[0.181 0.447]
$\psi_2 : \mathbb{E}_{t-1}y_t + \widehat{\Delta y}_{t,t-1}$	0.126	0.033	[0.063 0.191]
$\psi_1^* : (1-\rho_1-\rho_2)(\mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1})$	1.993	0.107	[1.772 2.200]
$\psi_2^* : (1-\rho_1-\rho_2)(\mathbb{E}_{t-1}\pi_t + \widehat{\Delta\pi}_{t,t-1})$	0.804	0.137	0.542 1.076]
σ^2	0.102	0.018	[0.072 0.143]
Log marginal likelihood: -16.95793			

Note: The table reports coefficient estimates for the specification (4.3) with $I = 2$. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters. The results indicate that all estimates are significant throughout the sample period. Models $\mathcal{M}_7$ - $\mathcal{M}_9$ correspond to estimates obtained using the OLS with real-time data, the one-step method with survey data, and the two-step method with survey data, respectively. Among these estimates, the one-step method yields the highest marginal likelihood. The rescaled parameters are defined as $\psi_1^* \equiv \psi_1/(1-\rho_1-\rho_2)$ and $\psi_2^* \equiv \psi_2/(1-\rho_1-\rho_2)$.

4.3 Fiscal Policy Rule Estimations Tables 6–8 compare OLS estimations using real-time data with the one-step and two-step methods using survey data across various specifications of government spending policy rules. Our findings demonstrate that the one-step and two-step methods yield higher marginal likelihoods than OLS estimations for all specifications incorporating policy inertia. Notably, the two-step method outperforms the one-step method in all government spending policy specifications with inertia. Although we cannot reject the null hypothesis of complete information from 1987 to 2006, this does not necessarily imply the absence of information rigidities.

Our results indicate that the specification with a one-period lag is statistically preferred. Specifically, when using real-time data, the coefficient on prior-period government spending is approximately 1.023. Moreover, the two-step method exhibits a higher log marginal likelihood in this specification compared to both the OLS estimation using real-time data and the one-step method using survey data. In contrast, the coefficients on the two-period lags of government spending are insignificant for both real-time and survey data, as the 2.5th and 97.5th percentiles for these coefficients span both positive and negative values. Furthermore, the two-step method achieves a higher log marginal likelihood in this specification than both the OLS estimation using real-time data and the one-step method using survey data.

Table 6: Estimates of fiscal policy without lag

$\mathcal{M}_1$	postmean	postsd	[0.025 0.975]
$\phi_1 : b_{t-1}$	0.424	0.009	[0.406 0.441]
$\phi_2 : y_t$	-0.004	0.002	[-0.007 -0.001]
σ^2	0.001	0.0002	[0.001 0.001]
Log marginal likelihood: 135.3074			
$\mathcal{M}_2$	postmean	postsd	[0.025 0.975]
$\phi_1 : b_{t-1}$	0.425	0.009	[0.407 0.443]
$\phi_2 : \mathbb{E}_{t-1}y_t$	-0.003	0.002	[-0.006 0.0004]
σ^2	0.001	0.0002	[0.001 0.001]
Log marginal likelihood: 134.36			
$\mathcal{M}_3$	postmean	postsd	[0.025 0.975]
$\phi_1 : b_{t-1}$	0.425	0.009	[0.407 0.443]
$\phi_2 : \mathbb{E}_{t-1}y_t + \widehat{\Delta y}_{t,t-1}$	-0.003	0.002	[-0.006 0.001]
σ^2	0.001	0.0002	[0.001 0.001]
Log marginal likelihood: 134.22			

Note: The table reports coefficient estimates for the specification (4.4) with $I = 0$. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters. The estimates of ϕ_1 exhibit statistical significance across all three estimation methods, whereas those of ϕ_2 are not. Models $\mathcal{M}_1$ - $\mathcal{M}_3$ correspond to estimates obtained using the OLS with real-time data, the one-step method with survey data, and the two-step method with survey data, respectively. Among these estimates, the one-step method yields the highest marginal likelihood.

Table 7: Estimates of fiscal policy with one-period lag

$\mathcal{M}_4$	postmean	postsd	[0.025 0.975]
$\varrho_1 : g_{t-1}$	1.023	0.006	[1.011 1.035]
$\phi_1 : b_{t-1}$	-0.010	0.003	[-0.016 -0.005]
$\phi_2 : y_t$	0.0001	0.0001	[-0.00003 0.0003]
$\phi_1^* : (1-\varrho_1)b_{t-1}$	0.457	0.029	[0.415 0.516]
$\phi_2^* : (1-\varrho_1)y_t$	-0.006	0.005	[-0.016 0.002]
σ^2	0.00000	0.00000	[0.00000 0.00000]
Log marginal likelihood: 329.5605			
$\mathcal{M}_5$	postmean	postsd	[0.025 0.975]
$\varrho_1 : g_{t-1}$	1.023	0.006	[1.012 1.035]
$\phi_1 : b_{t-1}$	-0.011	0.003	[-0.016 -0.005]
$\phi_2 : \mathbb{E}_{t-1}y_t$	0.0002	0.0001	[0.00002 0.0004]
$\phi_1^* : (1-\varrho_1)b_{t-1}$	0.452	0.025	[0.411 0.506]
$\phi_2^* : (1-\varrho_1)\mathbb{E}_{t-1}y_t$	-0.009	0.005	[-0.019 0.001]
σ^2	0.00000	0.00000	[0.00000 0.00000]
Log marginal likelihood: 330.7171			
$\mathcal{M}_6$	postmean	postsd	[0.025 0.975]
$\varrho_1 : g_{t-1}$	1.023	0.006	[1.012 1.035]
$\phi_1 : b_{t-1}$	-0.010	0.003	[-0.016 -0.005]
$\phi_2 : y_t$	0.0002	0.0001	[0.00003 0.0004]
$\phi_1^* : (1-\varrho_1)b_{t-1}$	0.453	0.025	[0.412 0.506]
$\phi_2^* : (1-\varrho_1)(\mathbb{E}_{t-1}y_t + \widehat{\Delta y}_{t,t-1})$	-0.010	0.005	[-0.020 0.001]
σ^2	0.00000	0.00000	[0.00000 0.00000]
Log marginal likelihood: 330.8498			

Note: The table reports coefficients estimates for the specification (4.4) with $I = 1$. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters. The results indicate that all estimates are significant throughout the sample period. The estimates of ϕ_1 and ϱ_1 exhibit statistical significance across all three estimation methods, whereas those of ϕ_2 and ϱ_2 are not. Models $\mathcal{M}_4$ - $\mathcal{M}_6$ correspond to estimates obtained using the OLS with real-time data, the one-step method with survey data, and the two-step method with survey data, respectively. Among these estimates, the one-step method yields the highest marginal likelihood. The rescaled parameters are defined as $\phi_1^* \equiv \phi_1/(1-\varrho_1)$ and $\phi_2^* \equiv \phi_2/(1-\varrho_1)$.

Table 8: Estimates of fiscal policy with two-period lag

$\mathcal{M}_7$	postmean	postsd	[0.025 0.975]
$\varrho_1 : g_{t-1}$	1.082	0.117	[0.852 1.313]
$\varrho_2 : g_{t-2}$	-0.061	0.120	[-0.296 0.175]
$\phi_1 : b_{t-1}$	-0.010	0.003	[-0.016 -0.004]
$\phi_2 : y_t$	0.0001	0.0001	[-0.00004 0.0003]
$\phi_1^* : (1-\varrho_1-\varrho_2)b_{t-1}$	0.459	0.058	[0.409 0.528]
$\phi_2^* : (1-\varrho_1-\varrho_2)y_t$	-0.007	0.009	[-0.019 0.002]
σ^2	0.00000	0.00000	[0.00000 0.00000]
Log marginal likelihood: 328.9756			
$\mathcal{M}_8$	postmean	postsd	[0.025 0.975]
$\varrho_1 : g_{t-1}$	1.081	0.115	[0.856 1.308]
$\varrho_2 : g_{t-2}$	-0.059	0.118	[-0.290 0.171]
$\phi_1 : b_{t-1}$	-0.010	0.003	[-0.015 -0.004]
$\phi_2 : y_t$	0.0002	0.0001	[0.00002 0.0004]
$\phi_1^* : (1-\varrho_1-\varrho_2)b_{t-1}$	0.453	0.091	[0.405 0.513]
$\phi_2^* : (1-\varrho_1-\varrho_2)\mathbb{E}_{t-1}y_t$	-0.010	0.037	[-0.023 -0.001]
σ^2	0.00000	0.00000	[0.00000 0.00000]
Log marginal likelihood: 330.1284			
$\mathcal{M}_9$	postmean	postsd	[0.025 0.975]
$\varrho_1 : g_{t-1}$	1.078	0.115	[0.853 1.304]
$\varrho_2 : g_{t-2}$	-0.056	0.118	[-0.286 0.174]
$\phi_1 : b_{t-1}$	-0.010	0.003	[-0.015 -0.004]
$\phi_2 : y_t$	0.0002	0.0001	[0.00003 0.0004]
$\phi_1^* : (1-\varrho_1-\varrho_2)b_{t-1}$	0.453	0.058	[0.406 0.513]
$\phi_2^* : (1-\varrho_1-\varrho_2)(\mathbb{E}_{t-1}y_t + \widehat{\Delta y}_{t,t-1})$	-0.010	0.023	[-0.023 -0.002]
σ^2	0.00000	0.00000	[0.00000 0.00000]
Log marginal likelihood: 330.2487			

Note: The table reports coefficients estimates for the specification (4.4) with $I = 2$. It reports the posterior mean, posterior standard deviation (postsd), and 2.5th and 97.5th percentiles of the estimated parameters. The results indicate that all estimates are significant throughout the sample period. Models $\mathcal{M}_7$ - $\mathcal{M}_9$ correspond to estimates obtained using the OLS with real-time data, the one-step method with survey data, and the two-step method with survey data, respectively. Among these estimates, the one-step method yields the highest marginal likelihood. The rescaled parameters are defined as $\phi_1^* \equiv \phi_1/(1-\varrho_1-\varrho_2)$ and $\phi_2^* \equiv \phi_2/(1-\varrho_1-\varrho_2)$.

5 Conclusion

This paper introduces empirical methods leveraging survey data to mitigate endogeneity bias in policy rule estimation across various information environments. We demonstrate that under complete information, our approach effectively eliminates endogeneity bias, while under incomplete information, we further orthogonalize forecast errors to address this bias. Our theoretical analysis establishes that these survey-based methods substantially reduce estimation bias in both information settings.

To quantify the magnitude of endogeneity bias, we first employ a standard new Keynesian model as our analytical framework. Our simulations reveal that conventional OLS methods can produce significantly inaccurate estimates, while incorporating survey data markedly improves estimation precision. We extend our analysis to a sticky information model, where simulations confirm that traditional OLS approaches yield biased outcomes. However, both our one-step and two-step methods substantially enhance estimation accuracy, with the two-step method further refining precision and bringing estimated policy coefficients closer to their true values.

We apply our proposed methodology to two key policy rules: interest rate and government spending rules. Our empirical results demonstrate that survey data significantly reduces endogeneity bias across both policy frameworks. Importantly, during our sample period, we cannot reject the null hypothesis of complete information, suggesting that information rigidities do not present a serious concern. This finding indicates that the one-step method is sufficient for our empirical purposes. Across nearly all model specifications, our survey-based approaches yield higher marginal likelihoods compared to standard OLS estimation, with the one-step method generally outperforming the two-step method in most specifications.

References

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2021): "Imperfect macroeconomic expectations: Evidence and theory," *NBER Macroeconomics Annual*, 35(1), 1–86.

Bunzel, H., and W. Enders (2010): "The Taylor rule and "opportunistic" monetary policy," *Journal of Money, Credit and Banking*, 42(5), 931–949.

Carvalho, C., F. Nechio, and T. Tristao (2021): "Taylor rule estimation by OLS," *Journal of Monetary Economics*, 124, 140–154.

Chib, S. (1995): "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313–1321.

Chib, S., and X. Zeng (2020): "Which factors are risk factors in asset pricing? A model scan framework," *Journal of Business & Economic Statistics*, 38(4), 771–783.

Coibion, O., and Y. Gorodnichenko (2012): "Why are target interest rate changes so persistent?," *American Economic Journal: Macroeconomics*, 4(4), 126–62.

——— (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *The American Economic Review*, 105(8), 2644–2678.

Galí, J. (2008): *Monetary Policy, Inflation, and the Business Cycle.* Princeton University Press, Princeton.

Orphanides, A. (2003): "Historical monetary policy analysis and the Taylor rule," *Journal of monetary economics*, 50(5), 983–1022.

Reis, R. (2009): "Optimal monetary policy rules in an estimated sticky-information model," *American Economic Journal: Macroeconomics*, 1(2), 1–28.

Taylor, J. B. (1993): "Discretion versus Policy Rules in Practice," *Carneige-Rochester Conference Series on Public Policy*, 39, 195–214.

Verona, F., and M. Wolters (2014): "Sticky Information Models in Dynare," *Comput Econ*, 43, 357–370.

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

——— (2003b): "Optimal Interest Rate Smoothing," *Review of Economic Studies*, 70(4), 861–886.

A Data description

Data are from National Income and Product Accounts (NIPA): the Bureau of Economic Analysis, Survey of Professional Forecasters (SPF), Tealbook (formerly Greenbook) Data Set, Federal Reserve Bank of Philadelphia, and Fed Funds Archival Economic Data, St. Louis Fed. Our main sample runs from 1987.Q3 to 2006.Q4 prior to the Great Recession, which is usually considered a structural break in the literature.

1. **Nominal interest rate**: nominal interest rates are Federal Funds Effective Rate (FEDFUNDS), obtained from Vintage Fed Funds Archival Economic Data, St. Louis Fed. Monthly data are extracted from vintages two quarters ahead. Quarterly data are averaged by monthly data. Vintage data start from 12/03/1996. Prior to 06/01/1996, data are extracted from the vintage of 12/03/1996. Starting from 07/01/1996, data are extracted from vintages two quarters ahead.

2. **Output Gap**: Real-time estimates and forecasts of output gaps are from an excel file named "Previous Output Format, Output Gap and Financial Assumptions". This data set contains real-time estimates and projections of the output gap used by the staff of the Board of Governors of the Federal Reserve System in constructing its Tealbook/Greenbook forecast.

   This dataset spans from the third quarter of 1987 through to the fourth quarter of 2013 and is part of the broader Tealbook Data Set, as of the fourth quarter of 2017. Formerly known as the Greenbook Data Set, this resource consolidates both real-time data and future forecasts concerning the output gap. The data set is particularly noteworthy for including estimates and projections that were instrumental for the Federal Reserve System's Board of Governors staff when formulating the Tealbook/Greenbook forecasts. Notably, these estimates and forecasts are based on information available two quarters ahead of their respective vintages, providing a forward-looking perspective on the output gap.

3. **Government spending**: Data are from National Income and Product Accounts (NIPA). Government consumption (Table 3.1, line 21) + government investment (Table 3.1, line 39) - consumption of fixed capital (Table 3.1, line 42).

4. **Inflation rate**: The Price Index for Gross National Product/Gross Domestic Product (PGDP), as maintained by the Federal Reserve Bank of Philadelphia, is a seasonally adjusted measure used to assess inflation. Real-time data are extracted from vintages two quarters ahead to maintain consistency with nominal interest rates. For data prior to June 1, 1996, values are taken from the December 3, 1996, vintage. Starting from July 1, 1996, data are extracted from vintages two quarters ahead. The inflation rate is computed as

the quarter-to-quarter logarithmic difference of PGDP and is annualized by multiplying the result by 400.

5. **Forecast of Inflation rate**: We use mean response data in SPF. PGDP1 is the current price index for GDP, while PGDP2 is the forecast of PGDP for the next quarter. The forecast of inflation rate is computed as the quarter-to-quarter logarithmic difference between PGDP2 and PGDP1. We convert the inflation rate into an annual percentage by multiplying by 400.