# Structural estimation of the New-Keynesian Model: a formal test of backward- and forward-looking expectations

Jang, Tae-Seok

University of Kiel

25 June 2012

# Structural estimation of the New-Keynesian model: a formal test of backward- and forward-looking expectations

Tae-Seok Jang*

[1] *Department of Economics, University of Kiel (CAU), Germany*

JUNE 25, 2012

**Abstract**

This paper attempts to uncover the empirical relationship between the price-setting/consumer behavior and the sources of persistence in inflation and output. First, a small-scale New-Keynesian model (NKM) is examined using the method of moment and maximum likelihood estimators with US data from 1960 to 2007. Then a formal test compares the fit of two competing specifications in the New-Keynesian Phillips Curve (NKPC) and the IS equation; i.e., forward- or backward-looking expectations. Accordingly, the inclusion of a lagged term in the NKPC and the IS equation improves the fit of the model while offsetting the influence of inherited and extrinsic persistence; it is shown that intrinsic persistence plays a major role in approximating the inflation and output dynamics for the *Great Inflation* period. However, the null hypothesis cannot be rejected at the 5% level for the *Great Moderation* period; i.e. the NKM of purely forward-looking behavior and its hybrid variant are equivalent. Monte Carlo experiments illustrate the validity of the chosen moment conditions and the finite sample properties of classical estimation methods. Finally, the empirical performance of model selection methods is investigated using the Akaike's and the Bayesian information criterion.

*JEL Classification: C12, C32, E12, G12*

*Keywords: formal test, foward- and backward-looking expectations, information criterion, intrinsic persistence, maximum likelihood, method of moment, New-Keynesian model*

*This paper was presented at the seminars in CAU and the 10th Annual *Advances in Econometrics* Conference at SMU. I have benefited from discussions with Reiner Franke, Ivan Jeliazkov, Vadim Marmer, Fabio Milani and Stephen Sacht. I also express my gratitude to all the participants for their active involvement in the conference. The support from the German Academic Exchange Service (DAAD) and the Möller fund from CAU is greatly acknowledged. Address of author: Department of Economics, University of Kiel, Olshausen Str. 40, D-24118, Kiel, Germany, E-Mail: taeseok.jang@gmail.com

# 1 Introduction

In the New-Keynesian model (NKM), some extensions such as the habit formation and indexing behavior have gained popularity for the ability to fit the macro data well; see Christiano et al. (2005), Smets and Wouters (2003, 2005, 2007) and Rabanal and Rubio-Ramarez (2005). For instance, the forward-looking behavior of price indexation has been challenged by macroeconomists over the last decade, because a hybrid variant of the model with the backward looking expectations provides a good approximation of inflation dynamics; see also Gali and Gertler (1999), Fuhrer (1997), Rudd and Whelan (2005, 2006), and many others. In the same way, inertial behavior in the dynamics of the output gap can be better explained by the presence of the habit formation in consumption rule; e.g. see Fuhrer (2000). Accordingly, the lagged dynamics in the NKM influence the transmission of shocks to the economy; the backward-looking behavior in the price-setting/consumption rules increases the degree of endogenous inflation/output persistence. This also implies that a good approximation of the NKM to the data (e.g. the persistence of aggregate macro variables) can provide a potential explanation for the monetary transmission channel to inflation and output; see Amato and Laubach (2003, 2004).

In a small-scale hybrid NKM, however, inflation and output depend on its expected future and lagged values, which induce a non-linear mapping between structural parameters and the objective function during estimation. Because of this, the structural system cannot avoid identification problems in the model; in other words, the minimization problem in extreme estimators often does not ensure a unique solution asymptotically; e.g. see Canova and Sala (2009). The purpose of this paper is to show to what extent classical estimation methods can cope with strucutural parameter estimates and how they can be used to evaluate the model's empirical performance. In particular, we focus on a system estimator which is based on an analytical solution of the model in estimation.[1]

More generally, to examine the significant influence of the lagged inflation and output on the structural dynamics, we apply the formal test of Hnatkovska, Marmer and Tang (2012) [HMT henceforth]. According to HMT, the Vuong-type $\chi^2$ test accomodates the adequacy of a broad class of the goodness-of-fit measures, which allow for model misspecifications; see also Linhart and Zucchini (1986) for model selection. Hence, the test statistic used herein can evaluate the discrepancy between the model-generated and empirical moments, which is associated with the goodness-of-fit of the model. With regard to the hypothesis testing, for example, the likelihood ratio test has been extensively used for non-nested hypotheses due to Vuong (1989). Rivers and Vuong (2002) generalized normal tests for model selection problems to the application involving a broad class of estimation methods. Their procedure extends to somewhat complex model selection situations where one or both models may be misspecified and the models may or may not be nested; see Golden (2000, 2003).

The advantage of the formal test of HMT is that the model's empirical performance can be flexibly evaluated according to the chosen moment conditions. The flexibilty is associated with the transparency to

[1] Alternatively, the common and simple strategy to provide a quantitative assement of inflation and output is to use a reduced form (or single equation) estimation, calibration or simulation based inference; see also Gregory and Smith (1991), Nason and Smith (2008) and others.

the fit of the model when the moment conditions are directly binding for parameter estimation. Indeed, the limited information approach has been widely used to estimate parameters of the DSGE models starting from Rotemberg and Woodford (1997). For instance, one common approach to this problem is to use impulse responses that are most informative about the DSGE models; Dridi et al. (2007) and Hall et al. (2011) discussed the choice of binding functions and information criteria for the selection of the valid response. Especially, when the model misspecifications and complex structural system do not allow for efficient estimation, the adequacy of the model in fitting the data can be judged by using binding functions; see Gourieroux and Monfort (1995). To provide parameter estimates using the limited information approach without auxiliary model, Franke et al. (2011) examined a small-scale DSGE model with analytical second moments of the sample auto- and cross-covariances up to lag 8 (two years) for estimation as well as model selection. While their empirical results are contrasted with the ones estimated by a Bayesian approach, however, the set of valid moment conditions is not tested.

In this paper, we discuss the efficiency of the method of moments (MM) estimation and examine the validity of moment conditions in comparison with the maximum likelihood (ML) approach. To see this, first, we study the relationship between interest rate, inflation and the output gap for US data, and provide a thorough analysis on empirical performance of the model and its connection to the expectation formation process: backward- and forward-looking behavior. From the ML and MM parameter estimates of the NKM, we pinpoint an empirical link between the hybrid model structure and the persistence in inflation and output. Next, the emprical performances of the NKM of purely forward and its hybrid variant are evaluated according to the model selection criterion. Accordingly, the inclusion of a lagged term in the New-Keynsian Phillips Curve (NKPC) and the IS equation improves the fit of the model while offsetting the influence of inherited and extrinsic persistence; it is shown that intrinsic persistence plays a major role in approximating the inflation and output dynamics for the *Great Inflation* period. However, the null hypothesis cannot be rejected at the 5% level for the *Great Moderation* period; i.e. the NKM of purely forward-looking behavior and its hybrid variant are equivalent. Finally, we also carry out a Monte Carlo (MC) study to examine the statistical efficiency of the estimation methods.

The paper is organized as follows: Section 2 reviews the standard New-Keynesian three-equations model and examines the importance of intrinsic persistence (or backward-looking behavior) for inflation and the output gap dynamics. Estimation methodologies and model selection procedures are described in section 3. Section 4 presents the empirical estimates of the competing models and discuss the results of their model comparison. The finite sample properties of MM and ML are investigated using the MC experiments in section 5. Finally, section 6 concludes. All technical details are collected in the appendix.

# 2 Expectation formation in a DSGE model

In this section, we present the standard New-Keynesian model featuring aggregate supply (AS), aggregate demand (IS), and monetary policy equations.[2] We focus on the model specifications for the expectation formation in the NKPC and the IS equation.

## 2.1 The New-Keynesian three-equations model

Microfoundations of supply- and demand-side economy have been established as the key components of a New-Keynesian model framework; e.g. the behavior of optimizing economic agents. The monetary policy behavior is described by the Taylor rule where the lagged interest rate reflects gradual adjustment of central banks. Thus the model is applicable to the dynamic analysis of economic changes. More generally, we attempt to examine the extent to which the gaps of interest rate, inflation and the output influence each other and affect the economy ($\widehat{\pi}_t := \pi_t - \pi_t^*,\ \widehat{r}_t := r_t - r_t^*$). The trend components of the quarterly data are estimated by using the Hodrick-Prescott filter with the smoothing parameter of $\lambda$=1600.[3]

$$
\begin{aligned}
\widehat{\pi}_t &= \frac{\beta}{1+\alpha\beta} E_t\, \widehat{\pi}_{t+1} + \frac{\alpha}{1+\alpha\beta} \widehat{\pi}_{t-1} + \kappa\, x_t + \nu_{\pi,t} \\
x_t &= \frac{1}{1+\chi} E_t\, x_{t+1} + \frac{\chi}{1+\chi} x_{t-1} - \tau\, (\widehat{r}_t - E_t\, \widehat{\pi}_{t+1}) + \nu_{x,t} \qquad (1) \\
\widehat{r}_t &= \phi_r\, \widehat{r}_{t-1} + (1-\phi_r)\, (\phi_\pi \widehat{\pi}_t + \phi_x x_t) + \varepsilon_{r,t}
\end{aligned}
$$

$$
\begin{aligned}
\nu_{\pi,t} &= \rho_\pi \nu_{\pi,t-1} + \varepsilon_{\pi,t} \quad \text{(for indexing behavior)} \qquad (2) \\
\nu_{x,t} &= \rho_x \nu_{x,t-1} + \varepsilon_{x,t} \quad \text{(for consumption behavior)}
\end{aligned}
$$

where the variable $x_t$ denotes the output gap, $\widehat{\pi}_t$ the inflation gap and $\widehat{r}_t$ the interest rate gap. The discount factor and the slope coefficient of the Phillips curve are denoted by the parameters $\beta$ and $\kappa$ respectively. The parameters $\alpha$ and $\chi$ measure the degree of price indexation in the NKPC ($0 \leq \alpha \leq 1$) and habit formation of the household ($0 \leq \chi \leq 1$) while $\tau$ relates to the intertemporal elasticity of substitution of consumption ($\tau \geq 0$). In the Taylor rule, $\phi_r$ determines the degree of interest rate smoothing ($0 \leq \phi_r \leq 1$). The other parameters $\phi_x$ and $\phi_\pi$ are the policy coefficients that measure the central bank's reactions to contemporaneous output and inflation ($\phi_x,\ \phi_\pi \geq 0$).

The shocks $\varepsilon_{z,t}$ are normally distributed with standard deviation $\sigma_z$ (i.i.d. with $z = \pi, x, r$). Since $\nu_{\pi,t}$ and $\nu_{x,t}$ are autoregressive processes, the persistences of the cost-push and the supply shocks are captured

[2]Smets and Wouters (2003, 2007) developed a medium-scale version of the NKM and estimated structural parameters and idiosyncratic shocks with the Bayesian techinques. In our study we focus on a small-scale general equilibrium model where we attempt to investigate the role of optimizing behavior in inflation and the output gap dynamics.

[3]Note here that we use the gaps instead of the levels for interest rate and inflation. Indeed, many empirical studies provide evidence for a time-varying trend in inflation and the natural rate of interest; see Castelnuovo (2010), Cogley and Sbordone (2008), Cogley et al. (2010) and many other studies. Further, the second moments are chosen to match the data when we estimate the model parameters. As a result, if we would use the non-stationary data without making assumptions about the DGP, it would cause substantial bias in parameter estimates of the structural model.

by the parameters $\rho_\pi$ and $\rho_x$, respectively ($0 \leq \rho_\pi, \rho_x \leq 1$). In estimation, we do not take them together, but treat them as being an independent case in order to directly disentangle the sources of inflation and the output gap persistence in the model.[4]

We denote $y_t$ by the state vector of three variables: $y_t = (\widehat{\pi}_t,\ x_t,\ \widehat{r}_t)'$. For the sake of simplicity, we present the above structural equations into the following canonical form:

$$\begin{aligned} AE_t y_{t+1} \;+\; By_t \;+\; Cy_{t-1} \;+\; \nu_t \;&= 0 \\ \nu_t \;=\; N\nu_{t-1} \;+\; \varepsilon_t, \qquad & \varepsilon_t \sim N(0, \Sigma_\varepsilon) \end{aligned} \tag{3}$$

With regard to the solution of the system, we begin with matrices $\Omega$ and $\Phi$ in the recursive equation of the reduced form. First, we use the method of undetermined coefficients to obtain the unique solution of the system under determinacy (i.e., $\phi_\pi \geq 1$). Second, we apply the brute force iteration method of Binder and Pesaran (1995) to numerically evaluate the matrix $\Omega$; see appendix B for some intermediate steps.

$$\begin{aligned} y_t &= \Omega y_{t-1} \;+\; \Phi \nu_t \\ \nu_t &= N\nu_{t-1} \;+\; \varepsilon_t \end{aligned} \tag{4}$$

From the matrices $\Omega$ and $\Phi$, it follows that the contemporaneous and lagged autocovariance process of the model can be computed recursively using the *Yule-Walker* equations; see chapter 2 of Lütkepohl (2005). On the whole, we only need to adjust the notation by changing the dating of the shocks and rewrite Equation (4) as

$$\begin{bmatrix} y_t \\ \nu_{t+1} \end{bmatrix} = \begin{bmatrix} \Omega & \Phi \\ 0 & N \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \nu_t \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \varepsilon_{t+1} \tag{5}$$

With $z_t = (y_t',\ \nu_{t+1}')'$, $D = (0\ I)'$, $u_t = D\varepsilon_{t+1}$, and $A_1$ the $2n \times 2n$ matrix on the righthand side associated with the vector $(y_{t-1}',\ \nu_t')' = z_{t-1}$, Equation (5) can be more compactly written as

$$z_t \;=\; A_1 z_{t-1} \;+\; u_t, \qquad u_t \sim N(0, \Sigma_u), \qquad \Sigma_u = D\Sigma_\varepsilon D' \tag{6}$$

where the matrix $A_1$ and the covariance matrix $\Sigma_u$ are functions of the parameter vector $\theta$. The estimation methodologies will be discussed later.

[4]The present study does not consider the presence of serially correlated shocks in the realizations of interest rate. It is assumed here that the shock persistence parameter of the interest rate gap ($\rho_r$) is explained by the lagged interest rate with its smoothing parameter ($\phi_r$). See also Carrillo et al. (2007) and many other studies.

## 2.2 Sources of persistence: backward- and forward-looking behavior

In the study of the model comparison, we put our focus on two polar cases of expectations. For example, when the price indexation parameter $\alpha$ is set to zero, it is assumed in the model that expectations are purely forward-looking. In this case, inflation persistence is exclusively driven by the exogenous shock process and inherited persistence from the output gap (see Table 1). But allowing it to be a free parameter, we assume that agents in the market form the backward-looking expectations. As a result, the NKPC is affected by both expected future and lagged inflation.

Table 1: Sources of persistence in the NKPC and the IS equation

| persistence | inflation | output gap |
|---|---|---|
| intrinsic | indexing behavior ($\alpha$) | habit formation ($\chi$) |
| extrinsic | AR (1) of the shock ($\rho_\pi$) | AR (1) of the shock ($\rho_x$) |
| inherited | slope of Phillips curve ($\kappa$) | intertemporal substitution ($\tau$) |

In the same vein, Table 1 shows that we can test the backward and forward-looking behavior in the IS equation. As long as each household chooses consumption optimally (i.e., without habit formation, $\chi = 0$), the persistence of the output gap is only driven by the exogenous shock and the inherited persistence implied by the Euler condition for the optimality intertemporal allocation of consumption. On the contrary, if habit formation is present in the consumption rule (i.e., $\chi$ is now a free parameter), then the output gap dynamics are endogeneously sustained by the optimizing behavior. As a result, the NKPC also depends on the lagged and expected future output gap.

In the present study, we aim to disentangle the sources of inflation and the output gap persistence using classical estimation methods. Note here that we capture each phenomenon by separately considering AR (1) of the shocks for the price indexing and consumption behavior.

# 3 Estimation methodologies and model selection

This section studies the structural estimation of the NKM when an analytical solution of the model is known. We also present the way in which a formal test can be used to evaluate the competing models.

## 3.1 Method of moment and model comparison: HMT (2012)

From the law of motion in Equation (6), it follows that the second moments of $z_t$ can be analytically computed. Thus the contemporaneous and lagged autocovariances of the first-order vector-autoregressive (VAR (1)) are given by:

$$\Gamma(h) \ := \ E(z_t z'_{t-h}) \in R^{K \times K}, \qquad K = 2n, \quad h = 0, 1, 2, \cdots \tag{7}$$

Their computation proceeds in two steps. First, $\Gamma(0)$ is obtained from the equation $\Gamma(0) = A_1 \Gamma(0) A'_1 + \Sigma_u$, which yields

$$\text{vec}\Gamma(0) \ = \ (I_{K^2} - A_1 \otimes A_1)^{-1} \text{vec}\Sigma_u \tag{8}$$

where the symbol '$\otimes$' denotes the Kronecker product and invertibility is guaranteed, because $A_1$ is clearly a stable matrix; i.e. $\phi_\pi \geq 1$. Subsequently the *Yule-Walker* equations are employed, from which we can recursively obtain the lagged autocovariances as

$$\Gamma(h) \ = \ A_1 \Gamma(h-1) \tag{9}$$

This formula relates to a vector autoregressive process of the model. From Equation (9), it follows that the model-generated second moments can be used to match the empirical counterparts in the MM estimation.

In order to compare the empirical performance of two models ($A$ and $B$), first we must estimate the model parameters by minimizing a weighted objective function (chosen goodness-of-fit measures):

$$J_I(\theta) \equiv \min_{\theta^I \in \Theta} \ \left\| W^{1/2} (\widehat{m}_T - m^I(\theta^I)) \right\|^2, \qquad I = A, B \tag{10}$$

where $m^I$ is a vector of moments and $\widehat{m}$ is a consistent and asymptotically normal estimator of true moments $m_0$. The norm of the matrix $X$ is defined as $||X|| = \sqrt{\text{tr}(X'X)}$, where tr denotes trace.

To examine the macroeconomic effects of the expected future and lagged dynamics in the NKPC and the IS equation, we use auto- and cross-covariances at lag 1 (15 moments) from the interest rate gap ($\widehat{r}_t$), the output gap ($x_t$), and the inflation rate gap ($\widehat{\pi}_t$); see also appendix A. As for the alternative moment conditions, we make use of the auto- and cross-covariances up to lag 4 (42 moments). The results of empirical findings and their robustness will be discussed later. Indeed, the second moments are applicable to the evaluation of the NKM's empirical performance and its comparison to the other specification.

In order to construct the objective function, we must estimate the weight matrix $W$ using the Newey-West estimator (Newey and West (1987))[5]:

$$\widehat{\Omega}_{NW} = \widehat{\Gamma}_T(0) + \sum_{k=1}^{5} \left( \widehat{\Gamma}_T(k) + \widehat{\Gamma}_T(k)' \right) \tag{11}$$

where $\widehat{\Gamma}_T(j)$ is $\frac{1}{T}\sum_{t=j+1}^{T}(m_t - \bar{m})(m_t - \bar{m})'$ and $k$ is the number of lags.[6] Note here that we use the diagonal components of the weight matrix and compute the inverse of $\widehat{\Omega}_{NW}$; here we impose the zero off-diagonal element restriction on the matrix $\widehat{\Omega}_{NW}$, because the elements of the weight matrix and the moments are highly correlated in a small sample size (see Altonji and Segal (1996)).

Under standard regularity conditions, the asymptotic distribution of the parameter estimates is given by:

$$\sqrt{T}(\widehat{\theta}_T - \theta_0) \sim N(0, \Lambda) \tag{12}$$

where we can numerically compute the covariance matrix $\Lambda$ using the first derivative of the moments at the optimum ($\Lambda = [(DWD')^{-1}]D'W\Omega WD[(DWD')^{-1}]'$).[7] Note here that $D$ is a gradient vector of moment functions evaluated at the estimated points:

$$\widehat{D}_T = \left. \frac{\partial m(\theta; X_T)}{\partial \theta} \right|_{\theta = \widehat{\theta}_T} \tag{13}$$

Next, we consider hypotheses comparing the goodness-of-fit of the competing models. The null hypothesis $H_0$ is that two non-nested models fit the data equally:

$$H_0 : \ \left\| W^{1/2}(\widehat{m}_T - m^A(\theta^A)) \right\| - \left\| W^{1/2}(\widehat{m}_T - m^B(\theta^B)) \right\| = 0 \tag{14}$$

[5] If large lags are included in the moments to be matched, the rows in the weight matrix are correlated to some extent. To avoid the dependence of the moments, we employ diagonal components of the Newey-West variance-covariance matrix in computing the weight matrix.

[6] The lag order is chosen following a simple rule of thumb for sample size ($\sim T^{1/4}$). For the GI and GM data, we have 78 and 99 quarterly observations respectively. Therfore $k$ is set to 5.

[7] If the weight matrix is chosen optimally ($\widehat{W} = \Omega^{-1}$), $\Lambda$ becomes $(DWD')^{-1}$; see chapter 1 of Anatolyev and Gospodinov (2011) among others. However, in our study, the estimated confidence bands become wider, because the weighting scheme in the objective function is not optimal.

The first alternative hypothesis is that model $A$ performs better than model $B$ when

$$H_1: \ \left\|W^{1/2}(\widehat{m}_T - m^A(\theta^A))\right\| - \left\|W^{1/2}(\widehat{m}_T - m^B(\theta^B))\right\| < 0 \tag{15}$$

The second alternative hypothesis is that model $B$ performs better than model $A$ when

$$H_2: \ \left\|W^{1/2}(\widehat{m}_T - m^A(\theta^A))\right\| - \left\|W^{1/2}(\widehat{m}_T - m^B(\theta^B))\right\| > 0 \tag{16}$$

To carry out the model comparison, we define the quasi-likelihood-ratio (QLR) statistic as

$$\widehat{\text{QLR}} = J^B(\widehat{\theta}^B) - J^A(\widehat{\theta}^A) \tag{17}$$

Following HMT, we consider the relationship between two models ($A$ and $B$): (i) nested, (ii) strictly non-nested and (iii) overlapping models. As long as the models share conditional distributions for the DGP and neither model is nested within the other, two models are overlapping. Then we can use a test of two sequential steps following Vuong (1989). To begin with, we compute critical values of the QLR distribution for the first step of the model comparison.[8] The simulated QLR distribution is defined as the following $\chi^2$-type formula:

$$Z' \ \widehat{\Sigma}_m^{1/2} W(V^B - V^A) W \ \widehat{\Sigma}_m^{1/2} Z, \qquad \text{where} \ \ Z \sim N(0, E_{n_m}) \tag{18}$$

where $\Sigma$ is a positive definite covariance matrix of the moment estimates and $Z$ is drawn from the multivariate ($n_m$) normal distribution. The $n_\theta^I$ by $n_\theta^I$ matrix $V^I$ is defined in appendix E. If $\widehat{\text{QLR}}$ exceeds the critical value from a 95% confidence interval, then the null hypothesis is rejected. Next, the second step tests whether or not the source of the rejection asymptotically comes from the same goodness-of-fit. The suggested test statistic has a standard normal distribution ($z$):

$$w_0 \ = \ 2 \ ||W_0^{1/2} \ (m^B(\theta^B) - m^A(\theta^A))|| \tag{19}$$

[8] Appendix E presents intermediate steps for simulating the QLR distribution. The theoretical QLR distribution is derived from the mean value expansion to a binding function (or moment conditions).

The standard deviation $w_0$ measures the uncertainty of the difference estimates between two models. Accordingly, the null of the equal fits can be rejected when $\sqrt{T} \cdot \mathrm{QLR}(\widehat{\theta}^B, \widehat{\theta}^A)/\widehat{w}_0 > z_{1-0.05/2}$ in which case $A$ is the preferred model, or $\sqrt{T} \cdot \mathrm{QLR}(\widehat{\theta}^B, \widehat{\theta}^A)/\widehat{w}_0 < -z_{1-0.05/2}$ in which case $B$ is preferred.

## 3.2 Maximum likelihood and model selection

The DSGE models have been examined using the ML estimation over the last decade; see Ireland (2004), Lindé (2005) and others. From the law of the motion in Equation (4), we may write that:

$$\begin{aligned} y_t &= \Omega y_{t-1} \ + \ \Phi \cdot (N \cdot \nu_{t-1} + \varepsilon_t) \\ &= (\Omega + \Phi N \Phi^{-1}) y_{t-1} \ - \ \Phi N \Phi^{-1} \Omega y_{t-2} \ + \ \Phi \cdot \varepsilon_t \end{aligned} \tag{20}$$

where we define the variable $\Phi \cdot \varepsilon_t$ as $\eta_t$. Now we assume that $\eta_t$ follows a multivariate normal distribution.

$$\eta_t \ \sim \ N(0, \Sigma_\eta), \qquad \Sigma_\eta \equiv \Phi \cdot \Sigma_\varepsilon \cdot \Phi' \tag{21}$$

Then we obtain the following conditional probability for the state variable $y_t$:

$$y_t | y_{t-1}, y_{t-2} \ \sim \ N((\Omega + \Phi N \Phi^{-1}) y_{t-1} - \Phi N \Phi^{-1} \Omega y_{t-2}, \ \Sigma_\eta) \tag{22}$$

Given the normality assumption of shocks and data set, the likelihood function can be constructed as:

$$L(\theta) = -\frac{n \cdot T}{2} \ln(2\pi) - \frac{T}{2} \ln |\Sigma_\eta| - \frac{1}{2} \sum_{t=2}^{T} \eta_t' \cdot \Sigma_\eta^{-1} \cdot \eta_t \tag{23}$$

where $n$ is the dimension of $y_t$. Then we arrive at the ML estimates for the parameter $\theta$ in optimization:

$$\theta_{ml} \ = \ arg \max_\theta \ L(\theta) \tag{24}$$

Under standard regularity conditions, the ML estimation is consistent and asymptotically normal:

$$\sqrt{T}(\widehat{\theta}_{ml} - \theta_0) \sim N(0, (\Upsilon/T)^{-1}) \tag{25}$$

where $\Upsilon = \mathrm{E}(\partial^2 L(\theta)/\partial\theta\partial\theta')$ is the information matrix. In our study, $\Upsilon$ is numerically computed using the Hessian matrix of the log likelihood function at optimum. To compare the likelihood values of the competing models, we use the well-known approach to model selection, the Akaike information criterion (AIC):

$$\mathrm{AIC} = -\frac{2}{T} \cdot \ln L(\theta) \;+\; \frac{2\,p}{T} \tag{26}$$

where $p$ is the dimension of the parameter $\theta$. Then, we choose the model for which AIC is the smallest. As an alternative to the AIC, which cannot respect the need for parsimony, we also consider the Bayesian information criterion (BIC):

$$\mathrm{BIC} = -\frac{2}{T} \cdot \ln L(\theta) \;+\; \frac{p \cdot \ln T}{T} \tag{27}$$

where the second term, $p \cdot \ln T$ penalizes the model with additional parameters.

# 4 Empirical application

In this section, we present the comparison of parameter estimates by MM and ML using the US data. First, we show how the persistence in inflaion and output can be disentangled in estimation. Second, we examine the empirical performance of the model using the formal test of HMT and discuss the similarities and dissimilarites between the MM and ML estiamtes. Finally, we also investigate the impact of the choice of moments on the parameter estimates.

## 4.1 Data

The data we use in this study comprise the GDP price deflator, the real GDP, and the federal funds rate. The series are taken from the US model datasets by Ray C. Fair; see the website (http://fairmodel.econ.yale.edu/main3.htm) for details. The trend rates underlying the gap formulation are treated as exogenously given. The trend from a Hodrick-Prescott (HP) filter is used with the smoothing parameter of $\lambda$ =1600. The data set covers the period 1960-2007. Due to the structural break beginning with the appointment of Paul Volcker as chairman of the U.S. Federal Reserve Board, we split data into two sub-samples: the *Great Inflation* (GI, 1960:Q1-1979:Q2) and the *Great Moderation* (GM, 1982:Q4-2007:Q2). The data split in the US economy is standard in much of the existing empirical work.

## 4.2 Basic results on method of moments estimation and model comparison

In this section, we use the MM estimation and attempt to estimate the parameters of two model specifications for inflation and the output gap persistence. Auto and cross-covariances at lag 1 are used as the chosen moment conditions; see appendix A. Next, we apply the model comparison method, which provides a formal assessment of the performance of competing specifications.

### 4.2.1 Assessing the fit of the model to inflation persistence: 15 moments

First, we examine the performance of the two models for fitting the GI data (see Table 2). As long as the profit maximizing rule (or purely forward-looking) determines the total amount of the output in the economy, the inflation dynamics are primarily captured by inherited and extrinsic persistence. Indeed, the model of purely forward-looking behavior has much higher estimated values for the parameters $\kappa$ and $\rho_\pi$ than its hybrid variant; i.e. $\widehat{\kappa}$: 0.12 (forward) > 0.05 (hybrid), $\widehat{\rho_\pi}$: 0.51 (forward) > 0.0 (hybrid).

Turning to the formal test, we classify the two models into the nested case. Since the hybrid variant of the model can generate richer dynamics due to the additional parameter $\alpha$, it nests the other model; the model with the forward-looking expectations does not allow the effects of lagged inflation on the NKPC.

To test the null hypothesis that the two models have an equal fit to the data, we compare the estimated loss function values ($\widehat{J}(\theta)$). We find QLR = 1.94. The simulated 1% and 5% critical values are 2.42 and 1.31, respectively; see the left panel of Figure 3 in appendix F. Therefore we reject the null hypothesis at the 5% level. This implies that the backward-looking behavior plays a significant role in approximating the inflation persistence of the GI.

Table 2: Parameter estimates for inflation persistence with 15 moments

| | GI | | GM | |
|---|---|---|---|---|
| | hybrid | forward | hybrid | forward |
| $\alpha$ | 0.768 | 0.0 (fixed) | 0.105 | 0.0 (fixed) |
| | (0.007 - 1.000) | ( - ) | (0.000 - 1.000) | ( - ) |
| $\kappa$ | 0.047 | 0.123 | 0.052 | 0.058 |
| | (0.009 - 0.084) | (0.000 - 0.318) | (0.000 - 0.136) | (0.008 - 0.107) |
| $\rho_\pi$ | 0.000 | 0.506 | 0.000 | 0.086 |
| | ( - ) | (0.078 - 0.933) | ( - ) | (0.000 - 0.269) |
| $\sigma_\pi$ | 0.679 | 0.778 | 0.638 | 0.644 |
| | (0.103 - 1.255) | (0.603 - 0.952) | (0.454 - 0.823) | (0.491 - 0.798) |
| $\chi$ | 1.000 | 0.999 | 0.774 | 0.802 |
| | ( - ) | (0.441 - 1.000) | (0.497 - 1.000) | (0.499 - 1.000) |
| $\tau$ | 0.094 | 0.089 | 0.000 | 0.000 |
| | (0.015 - 0.174) | (0.000 - 0.192) | ( - ) | ( - ) |
| $\rho_x$ | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) |
| | ( - ) | ( - ) | ( - ) | ( - ) |
| $\sigma_x$ | 0.727 | 0.662 | 0.404 | 0.369 |
| | (0.547 - 0.907) | (0.416 - 0.909) | (0.118 - 0.691) | (0.068 - 0.671) |
| $\phi_\pi$ | 1.659 | 1.744 | 1.798 | 1.943 |
| | (1.000 - 2.334) | (1.084 - 2.404) | (1.000 - 4.039) | (1.000 - 4.465) |
| $\phi_x$ | 0.378 | 0.181 | 0.729 | 0.652 |
| | (0.026 - 0.731) | (0.000 - 0.452) | (0.226 - 1.231) | (0.087 - 1.217) |
| $\phi_r$ | 0.544 | 0.463 | 0.841 | 0.849 |
| | (0.323 - 0.765) | (0.248 - 0.678) | (0.698 - 0.984) | (0.707 - 0.991) |
| $\sigma_r$ | 0.786 | 0.662 | 0.391 | 0.384 |
| | (0.382 - 1.190) | (0.155 - 1.169) | (0.099 - 0.684) | (0.080 - 0.688) |
| $J(\theta)$ | 1.30 | 3.24 | 2.26 | 2.44 |

Note: The discount factor parameter $\beta$ is calibrated to 0.99. The 95% asymptotic confidence intervals are given in brackets.

This finding is shown in Table 3. In particular, the results show that the hybrid variant of the model can approximate the inflation dynamics better than the other; e.g. see Cov$(r_t,\ x_{t-k})$, Cov$(x_t,\ \pi_{t-k})$, Cov$(\pi_t,\ r_{t-k})$. Nevertheless, the fit of the nested model is not so bad, because the estimated values of auto- and crosscovariances at lag 1 lie within the 95% confidence intervals of the empirical moments. Note here that we do not aim to match the auto- and cross-covariances up to higher lags; this will be discussed later.

Next, we take the same steps for the model comparison using the GM data. However, most parameter estimates of the two models do not differ too much. For instance, the estimated value for the price indexation is close to zero in the hybrid variant of the model; i.e. $\widehat{\alpha} = 0.105$. As a result, the result of the formal test shows that the two models fit the data equally well. First, we find QLR = 0.17. The simulated 1% and 5% criteria are 0.51 and 0.27, respectively; see the right panel of Figure 3 in appendix F. Therefore the null hypothesis cannot be rejected.

Table 3: Empirical and model-generated moments for inflation persistence: 15 moment conditions

| Label | Emp. | 95% CI | hybrid | forward | Label | Emp. | 95% CI | hybrid | forward |
|---|---|---|---|---|---|---|---|---|---|
| Var($r_t$) | 3.296 | 1.297-5.296 | 3.400 | 3.524 | Cov($x_t$,$x_{-1}$) | 2.523 | 1.356-3.690 | 2.365 | 2.495 |
| **Cov($r_t, r_{-1}$)** | 2.886 | 1.142-4.629 | **2.572** | 2.388 | **Cov($x_t, \pi_t$)** | 0.069 | -0.415-0.552 | **0.160** | 0.236 |
| **Cov($r_t, x_t$)** | 0.232 | -0.611-1.075 | **0.256** | 0.270 | **Cov($x_t, \pi_{-1}$)** | -0.350 | -1.239-0.539 | **-0.342** | -0.234 |
| **Cov($r_t, x_{-1}$)** | 0.991 | 0.235-1.746 | **0.946** | 0.782 | **Cov($\pi_t, r_{-1}$)** | 1.288 | -0.021-2.597 | **1.067** | 0.846 |
| **Cov($r_t, \pi_t$)** | 1.535 | -0.026-3.097 | **1.854** | 2.155 | **Cov($\pi_t, x_{-1}$)** | 0.588 | 0.199-0.977 | **0.527** | 0.442 |
| Cov($x_t, \pi_{-1}$) | 1.401 | 0.038-2.765 | 1.731 | 1.714 | Var($\pi_t$) | 1.989 | 0.615-3.364 | 1.713 | 1.921 |
| Cov($x_t, r_{-1}$) | -0.450 | -1.622-0.722 | -0.490 | -0.369 | Cov($\pi_t, \pi_{-1}$) | 0.893 | -0.216-2.001 | 1.033 | 0.789 |
| Var($x_t$) | 3.001 | 1.728-4.275 | 3.191 | 3.176 | | | | | |

Note: 95% CI means the 95% asymptotic confidence intervals for empirical moments.

To save space, we do not report the model-generated moments for GM. Indeed, when we compare trajectories of the model-generated moments (i.e. hybrid and forward), the model covariance profiles almost overlap with each other. The two models provide a good fit to auto- and cross-covarainces at the short lag. More ambitious attempts to take the model to data will be discussed using alternative moment conditions later, because the model has a bad fit to the ones up to relatively large lags (two or three years).

### 4.2.2 Assessing the fit of the model to the output gap persistence: 15 moments

The estimated parameters for the model with or without a habit formation are displayed in Table 4; in the purely forward-looking behavior, $\chi$ is set to zero, whereas this parameter is subject to the estimation in the hybrid variant of the model. The MM estimates of the two models have almost similar values except for the degree of the supply shock ($\sigma_x$) and the Taylor rule coefficient ($\phi_\pi$).

It can be seen from the GI data that the estimated value for the supply shocks with the model of an optimal consumer behavior is two times higher than the parameter value of the other model ($\widehat{\sigma}_x$: 0.45 (forward) $>$ 0.21 (hybrid)). This implies that the output gap dynamics are more or less driven by the high level of the supply shocks when a simple rule of thumb behavior is not allowed in the IS equation. As a result, the persistence from the supply shocks is transmitted to inflation, which is indicated by a lower value for the estimated price indexation parameter; i.e. $\widehat{\alpha}$: 0.517 (hybrid) $<$ 0.740 (forward). Further, concerning the model, which allows a fraction of consumers to have a rule of thumb behavior, the estimation results indicate a low value for the monetary coefficients on the inflation gap; i.e. $\widehat{\phi}_\pi$: 2.26 (forward) $>$ 1.86 (hybrid). Put differently, central banks react weakly to shocks due to the fact that the transmission of the shocks may endogenously affect the output gap persistence; since the parameter estimates are imprecise with a large confidence interval, however, this implication is not warranted. The reliability of the parameter estimates will be investigated later via a Monte Carlo study.

Table 4: Parameter estimates for the output gap persistence with 15 moments

| | GI | | GM | |
|---|---|---|---|---|
| | hybrid | forward | hybrid | forward |
| $\alpha$ | 0.517 | 0.740 | 0.039 | 0.036 |
| | (0.044 - 0.990) | (0.204 - 1.000) | (0.000 - 0.215) | (0.000 - 0.205) |
| $\kappa$ | 0.061 | 0.066 | 0.064 | 0.057 |
| | (0.011 - 0.112) | (0.004 - 0.128) | (0.000 - 0.130) | (0.000 - 0.117) |
| $\rho_\pi$ | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) |
| | ( - ) | ( - ) | ( - ) | ( - ) |
| $\sigma_\pi$ | 0.876 | 0.715 | 0.684 | 0.687 |
| | (0.576 - 1.175) | (0.447 - 0.983) | (0.545 - 0.824) | (0.547 - 0.826) |
| $\chi$ | 0.931 | 0.0 (fixed) | 0.585 | 0.0 (fixed) |
| | (0.000 - 1.000) | ( - ) | (0.000 - 1.000) | ( - ) |
| $\tau$ | 0.441 | 0.422 | 0.480 | 0.506 |
| | (0.000 - 0.943) | (0.000 - 0.995) | (0.000 - 1.223) | (0.000 - 1.315) |
| $\rho_x$ | 0.914 | 0.868 | 0.930 | 0.941 |
| | (0.756 - 1.000) | (0.725 - 1.000) | (0.864 - 0.996) | (0.878 - 1.000) |
| $\sigma_x$ | 0.214 | 0.445 | 0.197 | 0.218 |
| | (0.039 - 0.390) | (0.154 - 0.736) | (0.000 - 0.452) | (0.011 - 0.425) |
| $\phi_\pi$ | 1.857 | 2.256 | 1.109 | 1.354 |
| | (1.000 - 2.729) | (1.000 - 3.661) | (1.000 - 2.395) | (1.000 - 2.905) |
| $\phi_x$ | 0.838 | 0.797 | 1.526 | 1.438 |
| | (0.227 - 1.449) | (0.244 - 1.349) | (0.537 - 2.515) | (0.464 - 2.412) |
| $\phi_r$ | 0.725 | 0.835 | 0.863 | 0.898 |
| | (0.482 - 0.968) | (0.681 - 0.989) | (0.773 - 0.953) | (0.804 - 0.993) |
| $\sigma_r$ | 0.695 | 0.240 | 0.294 | 0.215 |
| | (0.207 - 1.183) | (0.000 - 1.326) | (0.060 - 0.528) | (0.000 - 0.612) |
| $J(\theta)$ | 0.44 | 1.91 | 0.40 | 0.57 |

Note: The discount factor parameter $\beta$ is calibrated to 0.99. The 95% asymptotic confidence intervals are given in brackets.

Now, we use the loss function values to provide a formal test to the two specifications in the IS equation. In GI, these values are respectively 0.44 and 1.91 for the model with and without habit formation. The simulated 1% and 5% test criteria are 1.89 and 1.08, respectively; see the left panel of Figure 4. Since the estimated value for QLR exceeds the criterion at the 5% level, we reject the null hypothesis that the two models are equivalent. This implies that the output gap dynamics are better approximated by the consumption behavior in a rule of thumb manner. This finding is shown in Table 5. For instance, the covariance profiles of $(r_t,\ x_{t-k})$, $(x_t,\ x_{t-k})$, and $(\pi_t,\ \pi_{t-k})$ are better captured by the hybrid variant of the model.

In the period of GM, the parameter estimates for the two models are found to be similar with each other. This implies that the difference in the loss function values is small (i.e., $\text{QLR} = 0.17$). The simulated

1% and 5% test criteria are 7.58 and 12.37, respectively; see the right panel of Figure 4 in appendix F. We cannot reject the null hypothesis that the two models are equivalent. To save space, we do not report the model-generalted moments for the GM period; the covariance profiles from the two models more or less overlap with each other.

Table 5: Empirical and model-generated moments for the output gap persistence: 15 moment conditions

| Label | Emp. | 95% CI | hybrid | forward | Label | Emp. | 95% CI | hybrid | forward |
|---|---|---|---|---|---|---|---|---|---|
| Var$(r_t)$ | 3.296 | 1.297 $\sim$ 5.296 | 3.305 | 3.196 | **Cov**$(x_t,x_{-1})$ | 2.523 | 1.356 $\sim$ 3.690 | **2.468** | 2.187 |
| **Cov**$(r_t, r_{-1})$ | 2.886 | 1.142 $\sim$ 4.629 | **2.873** | 3.041 | Cov$(x_t, \pi_t)$ | 0.069 | -0.415 $\sim$ 0.552 | 0.094 | 0.073 |
| Cov$(r_t, x_t)$ | 0.232 | -0.611 $\sim$ 1.075 | 0.164 | 0.342 | Cov$(x_t, \pi_{-1})$ | -0.350 | -1.239 $\sim$ 0.539 | -0.417 | -0.368 |
| **Cov**$(r_t, x_{-1})$ | 0.991 | 0.235 $\sim$ 1.746 | **0.984** | 0.789 | Cov$(\pi_t, r_{-1})$ | 1.288 | -0.021 $\sim$ 2.597 | 1.048 | 1.025 |
| Cov$(r_t, \pi_t)$ | 1.535 | -0.026 $\sim$ 3.097 | 1.657 | 1.525 | Cov$(\pi_t, x_{-1})$ | 0.588 | 0.199 $\sim$ 0.977 | 0.578 | 0.579 |
| Cov$(x_t, \pi_{-1})$ | 1.401 | 0.038 $\sim$ 2.765 | 1.582 | 1.638 | **Var**$(\pi_t)$ | 1.989 | 0.615 $\sim$ 3.364 | **1.907** | 1.810 |
| **Cov**$(x_t, r_{-1})$ | -0.450 | -1.622 $\sim$ 0.722 | **-0.252** | -0.073 | **Cov**$(\pi_t, \pi_{-1})$ | 0.893 | -0.216 $\sim$ 2.001 | **0.934** | 1.109 |
| **Var**$(x_t)$ | 3.001 | 1.728 $\sim$ 4.275 | **3.067** | 3.331 | | | | | |

Note: 95% CI means the 95% asymptotic confidence intervals for empirical moments.

### 4.3 Basic results on the maximum likelihood estimation

For comparison purposes, we present the ML estimates of the NKM, because it is often uncommon to see that the MM estimation includes all relevant information about the DGP; the MM estimation is likely to be as efficient as ML when the chosen moment conditions encompass as many features of the data as possible. Table 6 shows that ML and MM give somewhat similar parameter estimates to the hybrid variant of the model for inflation persistence. For instance, the parameter estimates for the price indexation $\alpha$ are 0.45 and 0.16 for the GI and GM data, respectively. The ML estimates also support the existence of intrinsic inflation persistence in the model. In other words, a rule of thumb behavior in the price-setting rule accounts for inflation persistence. Moreover, the ML estimation gives a very small value for the slope of the Phillips curve ($\widehat{\kappa} = 0.0$ (GI) and 0.04 (GM)). This implies that individual firms are less responsive to changes in economic activity (i.e., the Phillips curve is flat). Hence, inflation dynamics in GI are primarily driven by intrinsic (moderate) and extrinsic (strong) persistence; i.e. $\widehat{\alpha}$: 0.446, $\widehat{\sigma}_\pi$: 0.879.

In comparison, we find a slight difference for the estimation of the output gap persistence. For instance, the comparison of the estimation results between ML and MM shows that MM gives a much lower value for the habit formation parameter ($\chi$=0.28 and 0.25 for the GI and GM data). Further interesting observation from Table 6 is that the ML estimates for the intertemporal elasticity of substitution is found to be much lower ($\tau$=0.08 and 0.03 for the GI and GM data). This implies that intrinsic persistence in the output gap dynamics is less affected by the substitution effects implied by the Fisher equation.

The slight difference between the ML and MM estimates can be attributed to the the assumption of normality of the shocks; if the model is correctly specified, the ML estimates may be superior to the ones obtained by MM. Since we do not know the true DGP in almost all cases, however, MM may be a relevant choice for evaluating the model's goodnese-of-fit to the data; the moment matching results in a closer fit to the sample autocovariance. The statistical efficiency and consistency of the parameter estimation used in this study will be investigated via a Monte Carlo study later.

Table 6: ML estimates for inflation and the output gap persistence

| | Inflation Persistence | | | Output Gap Persistence | |
|---|---|---|---|---|---|
| | GI | GM | | GI | GM |
| $\alpha$ | 0.446 | 0.157 | $\alpha$ | 0.478 | 0.126 |
| | (0.241 - 0.652) | (0.149 - 0.164) | | (0.230 - 0.726) | (0.008 - 0.243) |
| $\kappa$ | 0.000 | 0.036 | $\kappa$ | 0.018 | 0.046 |
| | ( - ) | (0.034 - 0.037) | | (0.000 - 0.099) | (0.015 - 0.077) |
| $\rho_\pi$ | 0.000 | 0.000 | $\rho_\pi$ | 0.0 (fixed) | 0.0 (fixed) |
| | ( - ) | ( - ) | | ( - ) | ( - ) |
| $\sigma_\pi$ | 0.879 | 0.654 | $\sigma_\pi$ | 0.869 | 0.663 |
| | (0.740 - 1.019) | (0.649 - 0.660) | | (0.737 - 1.002) | (0.597 - 0.729) |
| $\chi$ | 1.000 | 0.998 | $\chi$ | 0.281 | 0.254 |
| | ( - ) | (0.978 - 1.000) | | (0.245 - 0.316) | (0.133 - 0.374) |
| $\tau$ | 0.037 | 0.016 | $\tau$ | 0.081 | 0.027 |
| | (0.001 - 0.073) | (0.014 - 0.019) | | (0.038 - 0.125) | (0.014 - 0.040) |
| $\rho_x$ | 0.0 (fixed) | 0.0 (fixed) | $\rho_x$ | 0.808 | 0.763 |
| | ( - ) | ( - ) | | (0.735 - 0.880) | (0.692 - 0.835) |
| $\sigma_x$ | 0.523 | 0.253 | $\sigma_x$ | 0.211 | 0.098 |
| | (0.442 - 0.604) | (0.252 - 0.255) | | (0.174 - 0.248) | (0.093 - 0.104) |
| $\phi_\pi$ | 1.353 | 1.001 | $\phi_\pi$ | 1.394 | 1.000 |
| | (1.000 - 2.760) | (1.000 - 1.112) | | (1.000 - 2.661) | ( - ) |
| $\phi_x$ | 1.180 | 1.275 | $\phi_x$ | 1.352 | 1.456 |
| | (0.295 - 2.064) | (1.225 - 1.324) | | (0.710 - 1.995) | (1.135 - 1.777) |
| $\phi_r$ | 0.809 | 0.830 | $\phi_r$ | 0.803 | 0.843 |
| | (0.690 - 0.927) | (0.827 - 0.833) | | (0.754 - 0.852) | (0.828 - 0.857) |
| $\sigma_r$ | 0.734 | 0.477 | $\sigma_r$ | 0.741 | 0.476 |
| | (0.618 - 0.850) | (0.472 - 0.481) | | (0.622 - 0.859) | (0.435 - 0.518) |
| $L(\theta)$ | -308.86 | -233.99 | $L(\theta)$ | -309.53 | -231.84 |

Note: The discount factor parameter $\beta$ is calibrated to 0.99. The 95% asymptotic confidence intervals are given in brackets.

Another point worth mentioning is that the high dimension of the parameter space can induce multiple local minima in the likelihood function. Once we change the starting values in optimization, we often obtain different values for the parameter estimates; more rigorous investigation with simulation-based optimization methods (i.e., simulated annealing, random search method) would be worthwhile. However, in the currrent study, we have a strong confidence in a unique global minimum for the parameter estimates, because we tested the parameter estimates with different starting values and found that they converge to the unique minimum points.

To make a more systemic investigation on our choice of moments in the model estimation, the next section examines the parameter estimates of the model using a large set of moment conditions.

## 4.4 Validity of extra moment conditions

In this section, we assess the sensitivity of the MM estimates to the changes in moment conditions. From this investigation, we will find that alternative moment conditions do not induce qualitative changes in the parameter estimation. To make our choice of moment condtions more reliable, first we present the vector autoregressive (VAR) model with lag 4 as a reference model; see appendix C for optimal lag selection criteria. Then we examine the persistence of the key macro data in the U.S. economy using auto- and cross-covariances up to lag 4.

### 4.4.1 Assessing the fit of the model to inflation persistence: 42 moments

With alternative moment conditions in hand (42 moments), we now estimate two specifications of the NKM: forward-looking ($\alpha = 0$) and hybrid case (i.e. $\alpha$ is a free parameter). Table 7 shows that the parameter estimates speak for strong backward-looking behavior in the NKM. And the MM estimates with a small and large set of moments give qualitatively similar values except for the policy shock parameter ($\sigma_r$=0.0). Indeed, ML would avoid such an estimate, given that there is a stochastic singularity with zero policy shock (i.e., the likelihood value becomes negative infinity at this point).

Next, we turn our attention to the model comparison. In the GI data, we found that the price indexation parameter is a corner solution. Accordingly we treat $\alpha$ as being exogenously fixed at unity, because HMT assume that the estimated parameters are in the interior of the admissible region (see their assumption 2.5 (b)). Put differently, since the price indexation parameter is set to different values, it can be seen that two models are now equally accurate and identical in population. In this respect, we treat two models as being overlapping and apply a two step sequential test for model comparison. On the contrary, a value for the estimated price indexation parameter lies in the interior of the parameter space for fitting the GM data ($\alpha = 0.525$). In this case, the hybrid version of the model nests the one with the purely forward-looking expectations.

In the period of GI, the hybrid variant of NKP has a better goodness-of-fit to the data ($J = 11.93$) than the purely forward-looking version of the model ($J = 42.77$). Indeed, the AR (1) coefficient for the cost push shock ($\rho_\pi$) plays an important role in the purely forward-looking NKP (see Table 7).[9] The results also show that inherited persistence has a significant impact on the output gap dynamics in the hybrid-variant of the model ($\widehat{\kappa} = 0.044$).

[9]The estimated value for the parameter $\sigma_r$ hit the boundary. This makes the objective function ill-behaved and partial derivatives numerically unstable. We set it to zero and compute the numerical derivatives of the other parameters for the model comparison. See appendix D for the matrix notation.

In order to examine the significant difference of moment estimates between the two specifications, we substract the objective function value of purely forward-looking NKM from the one of hybrid variant; i.e. QLR = 30.83. The difference exceeds the simulated criterion (95%) of the $\chi^2$-type distribution in the model comparison. According to the simulated test distribution, critical values for the 99% and 95% confidence intervals are 16.99 and 9.96 respectively (see the left panel of Figure 5 in appendix F). Since the test statistic exceeds the critical value at the 5% level, we proceed to take the second step of model comparison, which asymptotically examines the validity of the goodness-of-fit.

Table 7: Parameter estimates for inflation persistence with 42 moments

| | GI | | GM | |
|---|---|---|---|---|
| | hybrid | forward | hybrid | forward |
| $\alpha$ | 1.0 | 0.0 (fixed) | 0.509 | 0.0 (fixed) |
| | ( - ) | ( - ) | (0.126 - 0.924) | ( - ) |
| $\kappa$ | 0.044 | 0.155 | 0.037 | 0.102 |
| | (0.018 - 0.069) | (0.000 - 0.395) | (0.000 - 0.075) | (0.017 - 0.187) |
| $\rho_\pi$ | 0.000 | 0.675 | 0.000 | 0.596 |
| | ( - ) | (0.387 - 0.964) | (0.000 - 0.813) | (0.367 - 0.825) |
| $\sigma_\pi$ | 0.470 | 0.518 | 0.364 | 0.231 |
| | (0.000 - 1.686) | (0.233 - 0.790) | (0.048 - 0.680) | (0.093 - 0.369) |
| $\chi$ | 1.0 | 1.0 | 0.770 | 0.915 |
| | ( - ) | ( - ) | (0.515 - 1.000) | (0.518 - 1.000) |
| $\tau$ | 0.092 | 0.063 | 0.020 | 0.027 |
| | (0.045 - 0.140) | (0.008 - 0.118) | (0.000 - 0.055) | (0.000 - 0.074) |
| $\rho_x$ | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) |
| | ( - ) | ( - ) | ( - ) | ( - ) |
| $\sigma_x$ | 0.716 | 0.600 | 0.547 | 0.468 |
| | (0.462 - 0.970) | (0.348 - 0.853) | (0.202 - 0.820) | (0.185 - 0.751) |
| $\phi_\pi$ | 1.740 | 1.809 | 2.025 | 2.218 |
| | (1.255 - 2.225) | (1.221 - 2.397) | (1.000 - 2.870) | (1.141 - 3.114) |
| $\phi_x$ | 0.080 | 0.157 | 0.563 | 0.564 |
| | (0.000 - 0.542) | (0.000 - 0.528) | (0.216 - 1.059) | (0.154 - 0.974) |
| $\phi_r$ | 0.267 | 0.458 | 0.765 | 0.732 |
| | (0.000 - 0.905) | (0.224 - 0.692) | (0.619 - 0.881) | (0.592 - 0.872) |
| $\sigma_r$ | 0.000 | 0.000 | 0.486 | 0.545 |
| | ( - ) | ( - ) | (0.303 - 0.727) | (0.351 - 0.739) |
| $J(\theta)$ | 11.93 | 42.77 | 23.97 | 27.47 |

Note: The discount factor parameter $\beta$ is calibrated to 0.99. The 95% asymptotic confidence intervals are given in brackets.

In the second step of the formal test, we examine the uncertainty of the estimated difference between the two models for evaluating their fit to the data. We compute the plug-in estimate of $\widehat{w}_0$ (2.54). Under the null hypothesis, the test static follows a standard normal distribution; i.e. $\sqrt{T} \cdot QLR(\theta^A, \theta^B) \sim N(0, w_0^2)$. The estimate of $\sqrt{T} \cdot QLR/\widehat{w}$ is 1.37, which is smaller than a critical value at the 5% significance level of the two-tailed test. Therefore the results show that both models have the same goodness-of-fit and the null hypothesis cannot be rejected.[10] Figure 1 presents the model-generated moment conditions at three years for GI and contrasts them with the empirical estimates using a VAR (4) process.

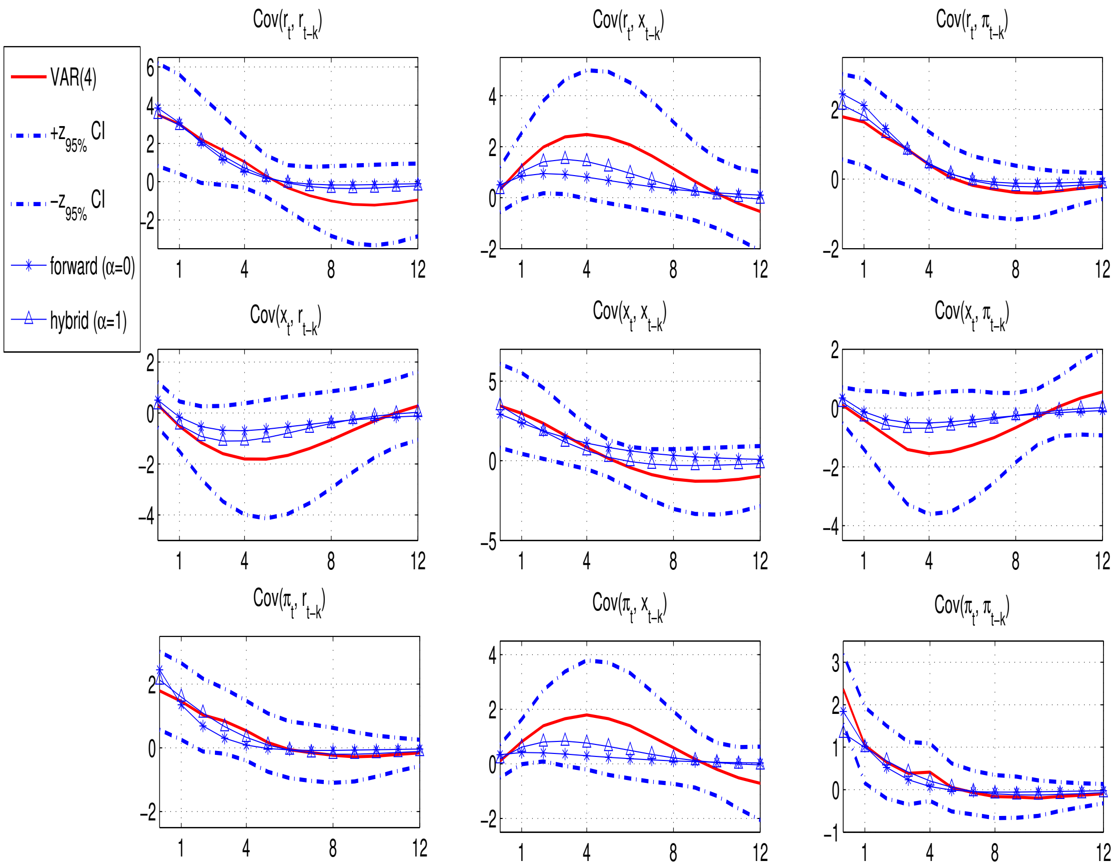


Figure 1: Covariance profiles for inflation persistence in GI (dashed: empirical, △: hybrid, *: forward)

Note: The empirical auto- and cross-covariances are computed using an unrestricted fourth-order vector autoregression (VAR) model. The asymptotic 95% confidence bands are constructed following Coenen (2005).

In the period of GM (Table 7), it is shown that the hybrid variant of NKP fits the data better (23.97). The estimation results indicate high values for the inherited and extrinsic persistence in the model of purely forward-looking behavior, because these can offset the impact of inherited persistence on the output gap dynamics; i.e. $\widehat{\kappa}$: 0.102 (forward) $>$ 0.037 (hybrid), $\widehat{\rho}_\pi$: 0.596 (forward) $>$ 0.0 (hybrid). However, the other parameter estimates are not different in both specifications.

[10] This statistical inference does not remain the same if the price indexation parameter is allowed to exceed unity. The contraint on habit formation parameter ($\chi$) is also removed. See Franke et al. (2011) for details.

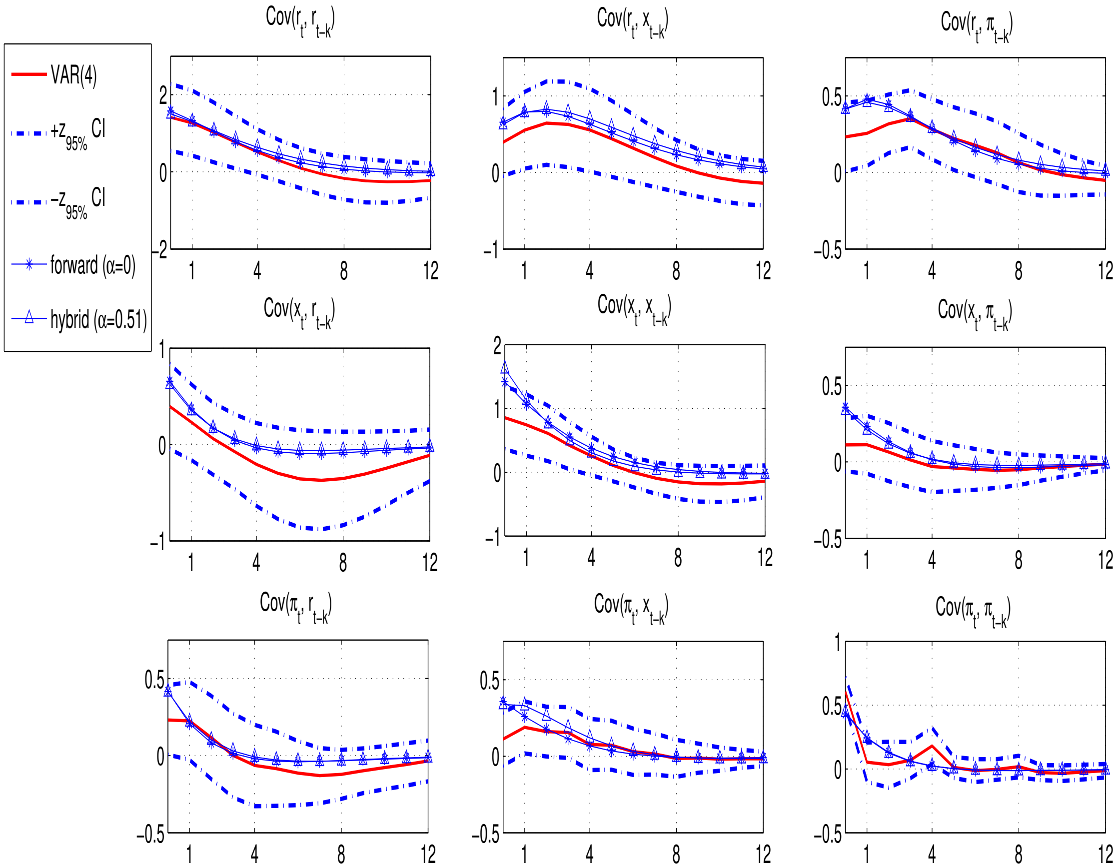


Figure 2: Covariance profiles for inflation persistence in GM (dashed: empirical, △: hybrid, *: forward)

Note: The empirical auto- and cross-covariances are computed using an unrestricted fourth-order vector autoregression (VAR) model. The asymptotic 95% confidence bands are constructed following Coenen (2005).

These empirical findings also seem to strengthen the relevance of backward-looking behaivior for the GM data. However, the difference between the two models (3.49) does not exceed the critical value for the 95% confidence intervals in the formal test; i.e., critical values for 99% and 95% confidence intervals are 38.39 and 21.46 respectively. Also see the right panel of Figure 5 in appendix F. Put differently, the effects of inherited persistence on the output gap can be adequately replaced by the inherited and extrinsic persistence, which cannot distinguish the sources of the persistence in the IS equation. Therefore we do not proceed to take the second step of the model comparison method and conclude that the null hypothesis cannot be rejected. Figure 2 presents the model-generated moment conditions at three years for the GM data; the comparison between the model-generated and the empirical moments by a VAR (4) process is displayed here.

### 4.4.2 Assessing the fit of the model to the output gap persistence: 42 moments

Table 8 shows the MM estimates for the output gap persistence using alternative moment conditions. Note here that the intertemporal elasticity of substitution of the both models has high estimated values in the GI and GM data; all the estimated values for $\rho_x$ exceed 0.7. In GI, this value increases substantially in the model with purely forward-looking expectations, which can cover the absence of intrinsic persistence in the IS equation; i.e. $\chi$=0.0 (fixed), $\widehat{\tau} = 0.676$.

Table 8: Parameter estimates for the output gap persistence with 42 moments

| | GI | | GM | |
|---|---|---|---|---|
| | hybrid | forward | hybrid | forward |
| $\alpha$ | 1.0 | 0.998 | 0.186 | 0.203 |
| | ( - ) | ( - ) | (0.000 - 0.396) | (0.000 - 0.441) |
| $\kappa$ | 0.054 | 0.037 | 0.086 | 0.088 |
| | (0.005 - 0.102) | (0.010 - 0.065) | (0.037 - 0.134) | (0.027 - 0.149) |
| $\rho_\pi$ | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) | 0.0 (fixed) |
| | ( - ) | ( - ) | ( - ) | ( - ) |
| $\sigma_\pi$ | 0.519 | 0.428 | 0.609 | 0.579 |
| | (0.099 - 0.939) | (0.108 - 0.747) | (0.461 - 0.757) | (0.410 - 0.749) |
| $\chi$ | 1.0 | 0.0 (fixed) | 0.991 | 0.0 (fixed) |
| | ( - ) | ( - ) | ( - ) | ( - ) |
| $\tau$ | 0.205 | 0.676 | 0.237 | 0.236 |
| | (0.000 - 0.436) | (0.000 - 1.897) | (0.000 - 0.547) | (0.000 - 0.803) |
| $\rho_x$ | 0.707 | 0.890 | 0.854 | 0.790 |
| | (0.290 - 1.000) | (0.743 - 1.000) | (0.686 - 1.000) | (0.583 - 0.997) |
| $\sigma_x$ | 0.213 | 0.519 | 0.140 | 0.340 |
| | (0.016 - 0.410) | (0.169 - 0.869) | (0.000 - 0.298) | (0.037 - 0.642) |
| $\phi_\pi$ | 1.741 | 2.046 | 2.133 | 2.224 |
| | (1.154 - 2.327) | (1.000 - 3.134) | (1.000 - 3.279) | (1.000 - 3.764) |
| $\phi_x$ | 0.169 | 1.103 | 0.762 | 0.588 |
| | (0.000 - 0.584) | (0.275 - 1.931) | (0.189 - 1.335) | (0.000 - 1.202) |
| $\phi_r$ | 0.389 | 0.889 | 0.770 | 0.783 |
| | (0.000 - 0.853) | (0.753 - 1.026) | (0.640 - 0.900) | (0.648 - 0.917) |
| $\sigma_r$ | 0.012 | 0.016 | 0.447 | 0.448 |
| | ( - ) | ( - ) | (0.248 - 0.645) | (0.212 - 0.685) |
| $J(\theta)$ | 10.54 | 31.64 | 20.79 | 23.85 |

Note: The discount factor parameter $\beta$ is calibrated to 0.99. The 95% asymptotic confidence intervals are given in brackets.

Another point worthwhile mentioning here is that the estimation results of the purely forward-looking model indicate high monetary policy coefficients on the interest rate gap, the inflation gap, the output gap. Further, in the hybrid variant, the parameter $\chi$ is almost a corner solution for both the GI and GM data, which strengthens a rule of thumb behavior in consumption. This implies that the rule of thumb behavior reinforces the degree of endogenous persistence in the output gap dynamics. However, as long as the model predicts that households behave optimally (i.e. without a simple rule of thumb behavior, $\chi = 0$), the result indicates the strong degree of the supply shocks; the estimated value is more than twice as high as the one of the hybrid model; i.e. $\widehat{\sigma}_x$: 0.519 (forward) $>$ 0.213 (hybrid) for GI, 0.340 (forward) $>$ 0.140 (hybrid) for GM).

As for the GI data, we treat the two models as being overlapping, because the habit formation parameter is now a corner solution. In the first step of the model comparison, we compare the objective function values (QLR = 21.10). The simulated 5% and 1% criteria are 19.63 and 34.59 respectively (see the left panel of Figure 6 in appendix F). Since the estimated QLR exceeds the 5% criterion value for the model comparison, we support the hypothesis that two models have different moments. In the second step, we estimate $\sqrt{T} \cdot QLR/\widehat{w}$ of which value is 1.02. However, this value does not exceed the criterion in the standard normal distribution. As a result, we conclude that there is no significant difference between two models in macthing the empirical moments; i.e. the two models have different moments, but an equivalent fit to the empirical moments. To save space, we do not provide the model covariance profiles for the output gap persistence. Note here that the result of the MM estimates with a large set of moments provides a closer fit (i.e. the sample auto- and cross-covariances up to large lags).

Now we turn our attention to the GM data. Note here that we treat the two model as being a nested case, since the estimated value for the habit formation parameter lies in an interior point. The model without the habit formation is nested within the other. First, we compute the difference between the objective function values of the two models (QLR = 3.06). Since the 5% and 1% criteria for the simulated test distribution are 18.52 and 29.05 respectively (see the right panel of Figure 7), however, the null hypothesis cannot be rejected. We conclude that two models have an equal fit to the empirical moments.

In sum, the MM estimates using a large set of moment conditions provide a stronger evidence for the backward-looking behavior in the price-setting and consumption rules compared to ML and MM with 15 moment conditions. This is a direct result when we include more sample second moments to be matched in the objective function. However, the formal model comparison test becomes inconclusive, because the estimated values for the price indexation and habit formation parameters were corner solutions; we used the two-step sequential hypothesis testing and found that the null hypothesis cannot be rejected, given that the sample size is small. The detailed discussions related to model selection will be discussed in the next section.

# 5 Attaining efficiency from moment conditions

In this section we study the finite sample properties of MM and ML; we investigate the role of model misspecification in the bias of the parameter estimation. Further, we discuss the empirical performance of model selection methods using the Akaike's and the Bayesian information criterion.

## 5.1 Monte Carlo study

The Monte Carlo (MC) experiment attempts to clearly demonstrate the statistical efficiency of the estimation methods, which are used in the previous section. Besides, we aim to investigate the role of the choice of moments and its influence on the parameter estimates. To begin, we consider the model specification of inflation persistence and set the parameters near to the values obtained by the MM estimation with 15 moments (see Table 2): e.g. high degree of backward-looking behavior ($\alpha$=0.750), moderate inherited persistence ($\kappa$=0.050), and no extrinsic persistence ($\rho_\pi$=0.0). Next, we generate 1,000 time series each consisting of 550 observations. The first 50 observations are removed as a transient period. Three sample sizes are considerd: 100, 200, and 500. We use the Matlab R2010a for this MC study As for the optimization routine, we use the unconstrained minimization "fminicon" with the algorithm 'interior-point'; maximum iteration and tolerance level are set to 500 and $10^{-6}$ respectively.

We conduct the MC experiments by considering two cases of model specification; i.e. correctly specified and misspecified. For the former, we discuss the finite sample properties of the MM and ML estimation. Turning to the misspecified case, we consider the model with purely forward-looking expecations and examine the degree of bias in the parameter estimates; i.e. (1) to what extent the extrinsic persistence ($\rho_\pi$) is inflated due to the misspecification and (2) how much the model misspecification affects the estimates for other strucutral parameters.

The main findings for the correctly specified case in Table 9 can be summarized as follows:

- For both ML and MM, the estimate of the price indexation paramter $\alpha$ is downward-biased, whereas the AR (1) coefficient of inflation shocks is estimated to be positive.
- ML has slightly poorer finite sample properties than MM. This implies that conventional Gaussian asymptotic approximation to the sample distribution is not as much precise as MM, as long as the sample size is small.
- The asymptotic efficiency of the ML estimates appears superior to MM, since the the mean of standard errors over 1000 estimations shows that the confidence intervals for the MM estimates are noticeably narrow. However, the large sample size remarkably improves the asymptotic efficiency of the MM estimates; e.g. T=500.
- It can be seen from the MC results that the overall parameter uncertainty of MM with a large set of moments is higher than ML and MM with a small set of moments. However, in this case, the MM estimation provides the most precise estimates on the price indexation parameter $\alpha$. Note here that the advantage of statistical inference for the backward-looking behavior comes at the cost of allowing

for large uncertainty in the estimates of other structural parameters; in other words, incorporating more second moments in the objective function improves the fit of the model to the persistence of inflation dynamics, but reduces efficiency in the estimates of other parameters.

- Another point, which is worthwhile to mention, is that we obtain the large asymptotic error for the policy shock parameter $\sigma_r$; S.E : 1.407 for T=100. This is attributed to the fact that the estimated values sometimes hit the boundary (i.e. $\sigma_r = 0.0$), which makes the numerical derivative unstable. This problem does not occur in the case where the large sample size is used (e.g. T=500).

With regard to the misspecified case, the MC results exhibit the high correlation between the price indexation and AR (1) coefficient of the inflation shocks; see appendix G. Indeed, it is shown in Table G.2 that the AR (1) coefficient is strongly upward-biased for both MM and ML, of which estimates offset the effects of intrinsic persistence on the inflation dynamics; e.g. $\rho_\pi$: 0.616 (ML), 0.632 (MM with 15 moments), 0.598 (MM with 42 moments) when the sample size is 100. The large sample size does not correct the bias of this parameter.

Similarly, the degree of the inflation shock $\sigma_\pi$ is more or less downward-biased. In addition, the slope coefficient of the Phillips curve is upward-biased in ML, and the result of the MM estimates shows very strong bias; $\widehat{\kappa}$: 0.096 (ML), 0.176 (MM with 15 moments), 0.205 (MM with 42 moments) when T=100. This implies that extrinsic (strong) and inherited (moderate) persistence offset the absence of intrinsic persistence due to the model misspecification. However, the other structural parameters are not influenced by the model misspecification; i.e. we obtain the parameter estimates near to the true ones by using both MM and ML. They converge at some reasonable rate towards the true parameters as the sample size gets larger (consistency).

Table 9: Monte Carlo results on the MM and ML estimates, ( ): root-mean-square-error, S.E : mean of standard error

| | | ML | | | MM with 15 moments | | | MM with 42 moments | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\theta^0$ | T = 100 | T = 200 | T = 500 | T = 100 | T = 200 | T = 500 | T = 100 | T = 200 | T = 500 |
| $\alpha$ | 0.750 | 0.523 (0.375) | 0.573 (0.322) | 0.651 (0.228) | 0.614 (0.256) | 0.654 (0.196) | 0.692 (0.121) | 0.700 (0.245) | 0.702 (0.205) | 0.729 (0.118) |
| | | S.E : 0.162 | S.E : 0.170 | S.E : 0.175 | S.E : 0.319 | S.E : 0.222 | S.E : 0.138 | S.E : 0.281 | S.E : 0.190 | S.E : 0.113 |
| $\kappa$ | 0.050 | 0.074 (0.076) | 0.066 (0.081) | 0.056 (0.014) | 0.083 (0.057) | 0.068 (0.030) | 0.058 (0.015) | 0.093 (0.075) | 0.073 (0.042) | 0.058 (0.018) |
| | | S.E : 0.054 | S.E : 0.048 | S.E : 0.041 | S.E : 0.042 | S.E : 0.025 | S.E : 0.013 | S.E : 0.050 | S.E : 0.030 | S.E : 0.014 |
| $\rho_\pi$ | 0.000 | 0.218 (0.330) | 0.172 (0.284) | 0.097 (0.198) | 0.175 (0.255) | 0.129 (0.194) | 0.082 (0.124) | 0.194 (0.299) | 0.147 (0.241) | 0.078 (0.144) |
| | | S.E : 0.112 | S.E : 0.1000 | S.E : 0.076 | S.E : 0.327 | S.E : 0.238 | S.E : 0.152 | S.E : 0.313 | S.E : 0.230 | S.E : 0.150 |
| $\sigma_\pi$ | 0.675 | 0.602 (0.330) | 0.619 (0.125) | 0.640 (0.073) | 0.613 (0.113) | 0.624 (0.085) | 0.639 (0.056) | 0.564 (0.1778) | 0.584 (0.136) | 0.618 (0.088) |
| | | S.E : 0.044 | S.E : 0.047 | S.E : 0.048 | S.E : 0.143 | S.E : 0.106 | S.E : 0.068 | S.E : 0.172 | S.E : 0.130 | S.E : 0.086 |
| $\chi$ | 1.000 | 0.935 (0.113) | 0.949 (0.090) | 0.967 (0.053) | 0.932 (0.108) | 0.948 (0.078) | 0.962 (0.055) | 0.941 (0.075) | 0.956 (0.083) | 0.966 (0.059) |
| | | S.E : 0.159 | S.E : 0.183 | S.E : 0.201 | S.E : 0.173 | S.E : 0.126 | S.E : 0.082 | S.E : 0.207 | S.E : 0.151 | S.E : 0.098 |
| $\tau$ | 0.090 | 0.089 (0.031) | 0.088 (0.023) | 0.087 (0.014) | 0.101 (0.039) | 0.095 (0.026) | 0.091 (0.016) | 0.105 (0.044) | 0.097 (0.030) | 0.092 (0.018) |
| | | S.E : 0.045 | S.E : 0.047 | S.E : 0.048 | S.E : 0.040 | S.E : 0.028 | S.E : 0.017 | S.E : 0.041 | S.E : 0.029 | S.E : 0.018 |
| $\sigma_x$ | 0.700 | 0.695 (0.059) | 0.697 (0.043) | 0.699 (0.025) | 0.743 (0.102) | 0.735 (0.073) | 0.724 (0.048) | 0.738 (0.123) | 0.729 (0.086) | 0.721 (0.054) |
| | | S.E : 0.050 | S.E : 0.052 | S.E : 0.053 | S.E : 0.086 | S.E : 0.062 | S.E : 0.039 | S.E : 0.121 | S.E : 0.089 | S.E : 0.057 |
| $\phi_\pi$ | 1.650 | 1.666 (0.183) | 1.654 (0.118) | 1.652 (0.074) | 1.681 (0.194) | 1.664 (0.123) | 1.659 (0.076) | 1.705 (0.229) | 1.679 (0.145) | 1.665 (0.088) |
| | | S.E : 0.345 | S.E : 0.316 | S.E : 0.274 | S.E : 0.210 | S.E : 0.147 | S.E : 0.093 | S.E : 0.214 | S.E : 0.151 | S.E : 0.098 |
| $\phi_x$ | 0.375 | 0.362 (0.124) | 0.361 (0.083) | 0.366 (0.052) | 0.337 (0.148) | 0.343 (0.100) | 0.352 (0.063) | 0.294 (0.191) | 0.317 (0.129) | 0.344 (0.082) |
| | | S.E : 0.227 | S.E : 0.224 | S.E : 0.228 | S.E : 0.137 | S.E : 0.097 | S.E : 0.062 | S.E : 0.156 | S.E : 0.110 | S.E : 0.071 |
| $\phi_r$ | 0.550 | 0.543 (0.048) | 0.545 (0.034) | 0.547 (0.021) | 0.525 (0.063) | 0.531 (0.045) | 0.538 (0.027) | 0.524 (0.080) | 0.532 (0.056) | 0.542 (0.034) |
| | | S.E : 0.068 | S.E : 0.070 | S.E : 0.077 | S.E : 0.074 | S.E : 0.052 | S.E : 0.033 | S.E : 0.086 | S.E : 0.061 | S.E : 0.039 |
| $\sigma_r$ | 0.750 | 0.738 (0.056) | 0.743 (0.038) | 0.748 (0.024) | 0.723 (0.087) | 0.736 (0.057) | 0.746 (0.034) | 0.617 (0.269) | 0.672 (0.173) | 0.721 (0.053) |
| | | S.E : 0.053 | S.E : 0.055 | S.E : 0.056 | S.E : 0.109 | S.E : 0.076 | S.E : 0.048 | S.E : 1.407 | S.E : 0.675 | S.E : 0.087 |
| $L(\theta)$ or $J(\theta)$ | | -385.76 | -800.93 | -2015.15 | 0.30 | 0.25 | 0.23 | 7.55 | 5.84 | 4.92 |

## 5.2 Model selection and discussion

From the empirical investigation on the MM estimates with a large set of moments, we found that the statistical power of the model comparison test is weak and the result becomes inconclusive; here we treat two models as being overlapping. Note here that we use the small sample to estimate the parameters of the NKM in which the asymptotic test of the model comparison is likely to make a Type II error; i.e. we accept the null hypothesis when the equal fit of moments is false.[11]

Table 10: Model selection using information criteria: inflation persistence

| | GI (T=78) | | | GM (T=99) | | |
|---|---|---|---|---|---|---|
| | ML | hybrid | forward | ML | hybrid | forward |
| $L(\theta)/T$ | -3.96 | -4.41 | -4.82 | -2.36 | -2.69 | -2.69 |
| AIC | 8.20 | 9.02 | 9.90 | 4.95 | 5.61 | 5.58 |
| BIC | 8.53 | 9.43 | 10.20 | 5.24 | 5.90 | 5.84 |
| Ranking | 1 | 2 | 3 | 1 | 3 | 2 |

Note: The backward- and forward-looking behaviors are examined using auto- and cross-covariances at lag 1.

To make the formal test more elaborate, we rank the model according to the well-known information criteria in the ML estimation. For this purpose, we suppose that the MM parameter estimates are possible solutions in the likelihood function. Table 10 and 11 report the mean value for the log-likelihood and the model selection criterion; the case of inflation and the output gap persistence respectively. Note here that we only present MM with a small set of the moment conditions (auto- and cross-covarainces at lag 1), because MM with alternative moments (auto- and cross-covarainces at lag 4) yields the zero policy shock for the GI data.

According to AIC and BIC, by definition, the ML estimates are preferred for both GI and GM data. If the assumption of normality is not violated and the model is correctly specified, we believe that the parameter estimates of the ML estimator are the most efficient; this statistical inference is verified by the MC study in the previous section. Nevertheless, the values for AIC and BIC using the MM estimation do not differ too much. This implies that matching the auto- and cross-covarainces at lag 1 provide more or less the same efficiency as the ML approach. Also the statistical inference for the expectation formation process does not change; i.e. the hybrid variant of the model can approximate inflation and the output gap dynamics better than the model with forward-looking behavior for fitting the GI data. On the other hand, the inconclusive result for the GM data shows that the model with forward-looking expectations of the price-setting rule is preferred due to its parsimonious description of the data.

[11] Marmer and Otsu (2012) studied the general optimality of comparison of misspecified models and proposed a feasible approximation to the optimal test which is more powerful than Rivers and Vuong (2002).

Table 11: Model selection using information criteria: the output gap persistence

| | GI (T=78) | | | GM (T=99) | | |
|---|---|---|---|---|---|---|
| | ML | hybrid | forward | ML | hybrid | forward |
| $L(\theta)/T$ | -3.97 | -4.62 | -7.88 | -2.34 | -3.09 | -4.22 |
| AIC | 8.22 | 9.51 | 16.01 | 4.91 | 6.41 | 8.64 |
| BIC | 8.55 | 9.85 | 16.31 | 5.19 | 6.69 | 8.90 |
| Ranking | 1 | 2 | 3 | 1 | 2 | 3 |

Note: The backward- and forward-looking behaviors are examined using auto- and cross-covariances at lag 1.

In brief, we can see from our empirical application that the moment-matching method achieves a high accuracy in taking the models to the data, but the parameter estimates are more uncertain than the ML estimates; i.e. wide confidence intervals. Indeed, this empirical observations can relate to the uncertainty of the model selection for expectation formation processes in the NKPC and the IS equation. In our empirical application, if we include additional second moments in the objective function, this improves the fit of the model to inflation and the output gap dynamics, but will make the comparison results of two models inconclusive. To address this issue on the trade-off between the fit of the model and the power of the formal test, we might clarify the model selection procedure in connection to the concept of model combination. The methods for the model averaging have been well developed in a Bayesian approach and will require another model selection framework to the applications involving the misspeicified restriction models.

# 6 Conclusion

This paper considered the structural estimation and model selection for expectation formation process in the NKM. We examined the importance of the future expected and lagged values in the inflation and output dynamics using US data; i.e. forward- and backward-looking expectations in the NKPC and the IS equation. The models are estimated by the classical estimation methods of MM and ML. In the former, we derived the analytical form of the auto- and cross-covariances in a linear system of the NKM; we estimate the parameters of the NKM by matching the model-generated moments with their empirical counterparts. These empirical findings are compared with the ones obtained by the ML estimation while their sensitivity to the moment conditions is also exmained.

According to the estimated loss function values obtained by MM, we evaluated two competing models using the formal test of HMT when they are overlapping or one model is nested within another. The results obtained with the GI and GM data show that the empirical performance of the NKM can be improved when allowing for backward-looking expectations in the NKP and the IS equation. After all, the backward-looking behavior in the NKM plays an important role in approximating inflation and the output gap dynamics in the DGP. This result suggests intrinsic persistence as the main source of inflation and the output gap dynamics in the GI data. However, we cannot reject the null hypothesis at the 5% level, because the model with purely forward-looking expecations and its hybrid-variant in the NKPC and IS equation have an equal fit to the GM data. These empirical findings are verified using the MC experiments; we investigated the statistical efficiency of the estimators and the implications for the model selection.

We close this paper by pointing out that (analytical) moment condtions provide information, which can be used to estimate structural parameters in the model; from this, we directly compare the competing specifications in the NKM using the formal test. Moreover, if the model does not have readily available expressions for moment conditions due to its non-linear model structure, they can be replaced by an approximation based on simulations. For example, De Grauwe (2010) developed a monetary DSGE model where agents' belief displays endogeneous waves of market optimism and pessimism. However, non-linear variants of DSGE models do not have a simple closed-form expression for a VAR (q) process. If this is the case, the simulated method of moments can be used to approximate the non-linearities in the model dynamics; see Jang and Sacht (2012) regarding simulation based inference for the non-linear group dynamics. Another example would be DSGE models with recursive preference and stochastic volatility (SV); i.e., see also Caldara et al. (2012) for the comparison of the solution methods. The non-linearity from recursive preferences and SV can be simply simulated and estimated via the method of moments used in this paper. We leave it to future research to empirically examine this kind of non-linear models.

## References


Altonji, J. and Segal, L. (1996): Small-sample bias in GMM estimation of covariance structures. *Journal of Business and Economic Statistics*, 14, 353–366.

Amato, J.D. and Laubach, T. (2003): Rule-of-thumb behaviour and monetary policy. *European Economic Review*, Vol. 47(5), pp. 791–831.

Amato, J.D. and Laubach, T. (2004): Implications of habit formation for optimal monetary policy. *Journal of Monetary Economics*, Vol. 51, pp. 305–325.

Anatolyev, S. and Gospodinov, N. (2011): *Methods for Estimation and Inference in Modern Econometrics.* New York: A Chapman and Hall Book.

Andrews, D. (1991), Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, Vol. 59, pp. 817–858.

Binder, M., and Pesaran, H. (1995): Multivariate Rational Expectations Models and Macroeconomic Modeling: a Review and Some Insights, In: Pesaran, M.H. and Wickens, M. (ed.), *Handbook of Applied Econometrics.* Oxford: Basil Blackwell, pp. 139–187.

Caldara, D., Fernandez-Villaverde, J., Rubio-Ramirez, J., and Yao, W. (2012): Computing DSGE models with recursive preferences and stochastic volatility. *Review of Economic Dynamics*, Vol. 15(2), pp. 188–206.

Canova, F. and Sala, L. (2009): Back to square one: identification issues in DSGE models. *Journal of Monetary Economics*, Vol. 56(4), pp. 431–449.

Carrillo, J.A., Féve, P. and Matheron J. (2007): Monetary policy inertia or persistent shocks: a DSGE analysis. *International Journal of Central Banking*, Vol. 3(2), pp. 1–38.

Castelnuovo, E. (2010): Trend inflation and macroeconomic volatilities in the Post WWII U.S. economy. *North American Journal of Economics and Finance*, Vol. 21, pp. 19–33.

Christiano, L., Eichenbaum, M. and C. Evans (2005): Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* Vol. 113(1), pp. 1–45.

Clarida, R., and Gali, J. and Gertler, M. (2000): Monetary policy rules and macroeconomic stability: evidence and some theory. *Quarterly Journal of Economics*, Vol. 115, pp. 147–180.

Coenen, G. (2005): Asymptotic confidence bands for the estimated autocovariance and autocorrelation functions of vector autoregressive models. *Empirical Economics*, Vol. 30(1), pp. 65–75.

Cogley, T., and Sbordone, A. (2008): Trend inflation, indexation, and inflation persistence in the New-Keynesian Phillips curve. *American Economic Review*, Vol. 98(5), pp. 2101–2126.

Cogley, T., Primiceri, G.E. and Sargent, T.J. (2010): Inflation-gap persistence in the US. *American Economic Journal: Macroeconomics*, Vol. 2(1), pp. 4369.

De Grauwe, P. (2010): Animal spirits and monetary policy. *Economic Theory*, Vol. 47, pp. 423–457.

Dridi, R. and Guay, A. and Renault, E. (2007): Indirect inference and calibration of dynamic stochastic general equilibrium models. *Journal of Econometrics*, Vol. 136(2), pp. 397-430.

Franke, R., and Jang, T.-S. and Sacht, S. (2011): *Moment matching versus Bayesian estimation: backward-looking behaviour in the New-Keynesian three-equations model.* Economic Working paper 2011-10, University of Kiel.

Fuhrer, J.C. (1997): The (un)importance of forward-Looking behavior in price specifications. *Journal of Money, Credit and Banking*, Vol. 29(3), pp. 338-350.

Fuhrer, J.C. (2000): Habit formation in consumption and its implications for monetary policy models. *American Economic Review*, Vol. 90(3), pp. 367-390.

Gali, J. and Gertler, M. (1999): Inflation dynamics: a structural econometric analysis. *Journal of Monetary Economics*, Vol. 44, pp. 195–222.

Gill, P. and Murray, W. and Wright, M. (1981): *Practical Optimization.* Academic Press.

Golden, R.M. (2000): Statistical tests for comparing possibly misspecified and nonnested models. *Journal of Mathematical Psychology*, Vol. 44, pp. 153–170.

Golden, R.M. (2003): Discrepancy risk model selection test theory for comparing possibly misspecified or nonnested models. *Psychometrika*, Vol. 68(2), pp. 229–249.

Gourieroux, C. and Monfort, A. (1995): Testing, encompassing and simulating dynamic econometric models. *Econometric Theory*, Vol. 11, pp. 195–228.

Gregory, A.W and Smith, G.W. (1991): Calibration as testing: inference in simulated macroeconomic models. *Journal of Business and Economic Statistics*, Vol. 9(3), pp. 297–303.

Hall, A.R., and Inoue, A., and Nason, J.M. and Rossi, B. (2011): Information criteria for impulse response function matching estimation of DSGE models, *Journal of Econometrics*, forthcoming.

Hnatkovska, V. and Marmer, V. and Tang, Y. (2012): Comparison of misspecified calibrated models: the minimum distance approach. *Journal of Econometrics, Forthcoming.*

Hnatkovska, V. and Marmer, V. and Tang, Y. (2011): *Supplement to "Comparison of misspecified calibrated models".* Working paper, University of Britisch Columbia.

Ireland, P.N. (2004): A method for taking models to the data. Journal of Economic Dynamics and Control, Vol. 28, pp. 1205–1226.

Jang, T.-S. and Sacht, S (2012): *Identification of animal spirits in a bounded rationality model: an application to the Euro area.* MPRA Paper No. 37399, University Library of Munich, Germany.

Lahiri, S.N. (2003): *Resampling Methods for Dependent Data.* Berlin: Springer.

Lee, B.-S. and Ingram, B.F. (1991): Simulation estimation of time-series models. *Journal of Econometrics*, Vol. 47, pp. 197-205.

Lindé, J. (2005): Estimating New-Keynesian Phillips curves: a full information maximum likelihood approach. *Journal of Monetary Economics*, Vol. 52, pp. 1135–49.

Linhart, H. and Zucchini, W. (1986): *Model Selection.* New York: John Wiley & Sons Inc.

Marmer, V. and Otsu, T. (2012): Optimal comparison of misspecified moment restriction models under a chosen measure of fit. *Journal of Econometrics, forthcoming.*

Nason, J.M. and Smith, G.W. (2008): Identifying the New Keynesian Phillips curve. *Journal of Applied Econometrics*, Vol. 23, pp. 525–551.

Newey, W. and West, K. (1987), A simple positive semi-definite, heteroskedasticty and autocorrelation consistent covariance matrix. *Econometrica*, Vol. 55, pp. 703–708.

Newey, W. and West, K. (1994): Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, Vol. 61(4), pp. 631–53.

RABANALA, P. AND RUBIO-RAMIREZB, J.F. (2005): Comparing New Keynesian models of the business cycle: a Bayesian approach. *Journal of Monetary Economics*, Vol. 52(6), pp. 1151–1166.

RIVERS, D. AND VUONG, Q. (2002): Model selection tests for nonlinear dynamic models. *Econometrics Journal*, Vol. 5, pp. 1–39.

ROTEMBERG, J. AND WOODFORD, M. (1997): An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual*, Cambridge, MA: MIT Press.

RUDD, J. AND WHELAN, K. (2005): New tests of the New-Keynsian Phillips curve. *Journal of Monetary Economics*, Vol. 52, pp. 1167–1181.

RUDD, J. AND WHELAN, K. (2006): Can rational expectations sticky-price models explain inflation dynamics. *American Economic Review*, Vol. 96(1), pp. 303–320.

SMETS, F. AND WOUTERS, R. (2003): An estimated stochastic dynamic general equilibrium model of the Euro area. *Journal of European Economic Association*, Vol. 1(5), pp. 1123–1175.

SMETS, F. AND WOUTERS, R. (2005): Comparing shocks and frictions in US and Euro area business cycles: a Bayesian DSGE approach. *Journal of Applied Econometrics*, Vol. 20(2), pp. 161–183.

SMETS, F. AND WOUTERS, R. (2007): Shocks and frictions in US business cycles: a Bayesian DSGE approach. *American Economic Review*, Vol. 97(3), pp. 586–606.

VUONG, Q. (1989): Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, Vol. 57(2), pp. 307–333.

# Appendices

# A Choice of moments

## A.1 Auto- and cross-covariances at lag 1 (one quarter): 15 moment conditions

This section lists the moment conditions for the method of moment estimation. The auto- and cross-covariances at lag 1 include the following 15 moment conditions after removing double counting of the interest gap ($\widehat{r}_t$), the ouput gap ($x_t$), and the inflation gap ($\widehat{\pi}_t$).

1. $m_1$: Var $(\widehat{r}_t)$
2. $m_2$: Cov $(\widehat{r}_t,\ \widehat{r}_{t-1})$
3. $m_3$: Cov $(\widehat{r}_t,\ x_t)$
4. $m_4$: Cov $(\widehat{r}_t,\ x_{t-1})$
5. $m_5$: Cov $(\widehat{r}_t,\ \widehat{\pi}_t)$
6. $m_6$: Cov $(\widehat{r}_t,\ \widehat{\pi}_{t-1})$
7. $m_7$: Cov $(x_t,\ \widehat{r}_{t-1})$
8. $m_8$: Var $(x_t)$
9. $m_9$: Cov $(x_t,\ x_{t-1})$
10. $m_{10}$: Cov $(x_t,\ \widehat{\pi}_t)$
11. $m_{11}$: Cov $(x_t,\ \widehat{\pi}_{t-1})$
12. $m_{12}$: Cov $(\widehat{\pi}_t,\ x_{t-1})$
13. $m_{13}$: Cov $(\widehat{\pi}_t,\ \widehat{r}_{t-1})$
14. $m_{14}$: Var $(\widehat{\pi}_t)$
15. $m_{15}$: Cov $(\widehat{\pi}_t,\ \widehat{\pi}_{t-1})$

## A.2 Auto- and cross-covariances at lag 4 (one year): 42 moment conditions

In the same vein, there are nine profies of the sample covariance functions. Counting all the combination of three state variables gives 42 moment conditions for the auto- and cross-covariances at lag 4. To save space, we abstract its list here by using the following notation:

$$\mathrm{Cov}(u_t,\ v_{t-h}), \qquad u\ \&\ v\ =\ \widehat{r}_t,\ x_t,\ \widehat{\pi}_t \tag{A.1}$$

where h denotes the lag length used in the auto- and cross-covarainces ($h = 0, 1, 2, 3, 4$).

# B Reduced form of matrix and solution of the NKM

In this section we give a description of the matrix notation and the solution procedure for the system of the NKM. The matrices of $A$, $B$, $C$ and $N$ in Equation (3) are defined as follows.

$$A = \begin{bmatrix} 0 & 0 & \frac{\beta}{1+\alpha\beta} \\ 0 & \frac{1}{1+\chi} & \tau \\ 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & \kappa & -1 \\ -\tau & -1 & 0 \\ -1 & (1-\phi_r)\phi_x & (1-\phi_r)\phi_\pi \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & \frac{\alpha}{1+\alpha\beta} \\ 0 & \frac{\chi}{1+\chi} & 0 \\ \phi_r & 0 & 0 \end{bmatrix}, \qquad N = \begin{bmatrix} 0 & 0 & \rho_\pi \\ 0 & \rho_x & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Using Equation (4), we redefine the state variable $y_t$ as terms of one-period-ahead.

$$\begin{aligned} y_{t+1} &= \Omega y_t + \Phi\nu_{t+1} \\ &= \Omega(\Omega y_{t-1} + \Phi\nu_t) + \Phi(N\nu_t + \varepsilon_{t+1}) \\ &= \Omega^2 y_{t-1} + (\Omega\Phi + \Phi N)\nu_t + \Phi\varepsilon_{t+1} \end{aligned} \tag{B.2}$$

Substitute Equations (B.2) and (4) into the canonical form of Equation (3).

$$E_t\big[A\Omega^2 y_{t-1} + A(\Omega\Phi + \Phi N)\nu_t + A\Phi\varepsilon_{t+1} + B\Omega y_{t-1} + B\Phi\nu_t + Cy_{t-1} + \nu_t\big] = 0 \tag{B.3}$$

Drop the expectation and rearrange things.

$$(A\Omega^2 + B\Omega + C)y_{t-1} + (A\Omega\Phi + A\Phi N + B\Phi + I_n)\nu_t = 0, \qquad \text{where } n = 3 \tag{B.4}$$

This implies that the following equations must hold for all $y_{t-1}$ and $\nu_t$.

$$\begin{aligned} A\Omega^2 + B\Omega + C &= 0 \\ (A\Omega + B)\Phi + A\Phi N + I_n &= 0 \end{aligned} \tag{B.5}$$

An iterative method can provide the solution of the matrix $\Omega$. The matrix $\Phi$ can be obtained by using some matrix algebra; i.e. the solution of the Lyapunov equation.

# C VAR lag order selection

A VAR (q) model has been used to estimate the empirical auto- and cross-covariances of interest, inflation and the output gap data. We use the model of a $K$-dimensional multiple times series $y_t := (y_{1t}, \cdots, y_{Kt})'$ following Lütkepohl (2005):

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_q y_{t-1} + u_t \tag{C.6}$$

where $\nu$ is a fixed $(K \times 1)$ vector of intercept and $u_t$ is a $K$-dimensional innovation process with $E(u_t) = 0$, $E(u_t u_t') = \Sigma_u$. The matrices $A_i$ include fixed $(K \times K)$ coefficients. The following lag order selection criteria are considered: final prediction error (FPE), Akaike information criterion (AIC), Hannan-Quinn information criterion (HQ), Bayesian information criterion (BIC). The chosen lag orders for both periods are one year (VAR (4)).

Table C.1: VAR lag order selection criteria

| | GI | | | | GM | | | |
|---|---|---|---|---|---|---|---|---|
| Lag | FPE | AIC | HQ | BIC | FPE | AIC | HQ | BIC |
| 0 | 14931.714 | 9.534 | 9.534 | 9.534 | 8926.601 | 9.036 | 9.036 | 9.036 |
| 1 | 194.525 | 5.309 | 5.302 | 5.466 | 205.437 | 5.554 | 5.558 | 5.699 |
| 2 | 106.200 | 4.822 | 4.805 | 5.137 | 112.227 | 4.843 | 4.851 | 5.136 |
| 3 | 24.202 | 3.462 | 3.435 | 3.936 | 26.806 | 3.505 | 3.515 | 3.945 |
| 4 | 1.136 | **0.522*** | **0.482*** | **1.156*** | 1.696 | **0.839*** | **0.851*** | **1.427*** |
| 5 | 1.058 | 0.569 | 0.515 | 1.365 | 1.759 | 0.970 | 0.983 | 1.708 |
| 6 | **0.944*** | 0.571 | 0.501 | 1.528 | 2.094 | 1.238 | 1.251 | 2.127 |
| 7 | 0.970 | 0.709 | 0.620 | 1.830 | 1.611 | 1.068 | 1.081 | 2.110 |
| 8 | 1.050 | 0.893 | 0.783 | 2.177 | **1.563*** | 1.129 | 1.139 | 2.324 |

Note: The star (*) indicates an optimal lag length.

# D Matrix notation

This section gives a matrix notation for the derivative of the moment conditions. This notation is used to implement the procedures for the model comparison of HMT; see appendix E. Let $m(\theta)$ be a $m_n$ by 1 vector. The parameter vector $\theta$ has a dimension of $n_\theta^I$. The gradient matrix $\frac{\partial m(\theta)}{\partial \theta'}$ has dimension $m_{m_n} \times n_\theta^I$. The second derivative matrix $\frac{\partial}{\partial \theta^{I\prime}} \text{vec}\big(\frac{\partial m^I(\theta^I)}{\partial \theta^{I\prime}}\big)$ has dimension $m_{m_n} \cdot n_\theta^I \times n_\theta^I$

$$\frac{\partial m(\theta)}{\partial \theta'} = \begin{bmatrix} \frac{\partial m_1}{\partial \theta_1} & \frac{\partial m_1}{\partial \theta_2} & \cdots & \frac{\partial m_1}{\partial \theta_{n_\theta^I}} \\ \frac{\partial m_2}{\partial \theta_1} & \frac{\partial m_2}{\partial \theta_2} & \cdots & \frac{\partial m_2}{\partial \theta_{n_\theta^I}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial m_{m_n}}{\partial \theta_1} & \frac{\partial m_{m_n}}{\partial \theta_2} & \cdots & \frac{\partial m_{m_n}}{\partial \theta_{n_\theta^I}} \end{bmatrix}.$$

$$\frac{\partial}{\partial \theta^{I\prime}} \text{vec}\left(\frac{\partial m^I(\theta^I)}{\partial \theta^{I\prime}}\right) = \begin{bmatrix} \frac{\partial m_1}{\partial \theta_1 \partial \theta_1} & \frac{\partial m_1}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial m_1}{\partial \theta_1 \partial \theta_{n_\theta^I}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial m_{m_n}}{\partial \theta_1 \partial \theta_1} & \frac{\partial m_{m_n}}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial m_{m_n}}{\partial \theta_1 \partial \theta_{n_\theta^I}} \\ \frac{\partial m_1}{\partial \theta_2 \partial \theta_1} & \frac{\partial m_1}{\partial \theta_2 \partial \theta_2} & \cdots & \frac{\partial m_1}{\partial \theta_2 \partial \theta_{n_\theta^I}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial m_{m_n}}{\partial \theta_2 \partial \theta_1} & \frac{\partial m_{m_n}}{\partial \theta_2 \partial \theta_2} & \cdots & \frac{\partial m_{m_n}}{\partial \theta_2 \partial \theta_{n_\theta^I}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial m_1}{\partial \theta_{n_\theta^I} \partial \theta_1} & \frac{\partial m_1}{\partial \theta_{n_\theta^I} \partial \theta_2} & \cdots & \frac{\partial m_1}{\partial \theta_{n_\theta^I} \partial \theta_{n_\theta^I}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial m_{m_n}}{\partial \theta_{n_\theta^I} \partial \theta_1} & \frac{\partial m_{m_n}}{\partial \theta_{n_\theta^I} \partial \theta_2} & \cdots & \frac{\partial m_{m_n}}{\partial \theta_{n_\theta^I} \partial \theta_{n_\theta^I}} \end{bmatrix}.$$

# E Technical note on the model comparison method

This section recapitulates the equations for the model comparision method of HMT. Assume that model $B$ is nested within model $A$. The quantitative goodness-of-fit of models to data is evaluated by using the method of moments in section 3.1. The "full" model is tested against the "restricted" model.

Let $m_T$ be a $n_m$ vector of moments. $\widehat{m}(\theta)$ is the consistent estimator of $m_T$. The uncertainty of moment estimates is assessed by estimating a Newey-West type weighted sum of autocovariance matrices ($\widehat{\Sigma}_m$). Given the assumption of normality, we can consistently estimate the covariance matrix of moment conditions.

$$\sqrt{T}(m_T - \widehat{m}(\theta)) \underset{d}{\rightarrow} N(0, \widehat{\Sigma}_m) \tag{E.7}$$

The estimates $\widehat{\theta}^I$ are obtained at the point where a weighted objective function is minimized:

$$J(\theta^I) \equiv \min_{\theta^I \in \Theta} \left\| W^{1/2}(\widehat{m}_T - m^I(\widehat{\theta^I}))\right\|^2, \qquad I = A, B \tag{E.8}$$

$\|W^{1/2}(\widehat{m}_T - m^I(\widehat{\theta^I}))\|$ is defined as $\sqrt{(\widehat{m}_T - m^I(\widehat{\theta}))' W (\widehat{m}_T - m^I(\widehat{\theta}))}$. The weight matrix $W$ is set to the diagonal components of $1/\hat{\Sigma}_{m,ii}$ ($ii = 1, \cdots, n_m$). The quasi-likelihood ratio test statistic is constructed as the difference in fits between two models:

$$\mathrm{QLR}(\widehat{\theta}^B, \widehat{\theta}^A) = J^B(\widehat{\theta}^B) - J^A(\widehat{\theta}^A) \tag{E.9}$$

$J^I$ ($I = A, B$) is a minimum value of the objective function given parameter estimates from Equation (E.8). It is assumed that the chosen moment functions in the models are twice continuously differentiable in neighborhoods of $\theta^I \subset \Theta^{n^I_\theta}$. Further, the matrix $F$ and $M$ are non-singluar in neighborhoods of $\theta$.[12]:

$$\begin{aligned} F^I &= \frac{\partial m^I(\theta^I)'}{\partial \theta^I} W \frac{\partial m^I(\theta^I)}{\partial \theta^{I'}} - M^I && \text{(E.10)} \\ M^I &= (E_I \otimes (\widehat{m}_T - m^I(\theta^I))'W)\frac{\partial}{\partial \theta^{I'}}\mathrm{vec}(\frac{\partial m^I(\theta^I)}{\partial \theta^{I'}}), \qquad I = A, B && \text{(E.11)} \end{aligned}$$

$E_I$ is the identity matrix of which dimension is $n^I_\theta \times n^I_\theta$. Note here that the dimensions of the matrices $\frac{\partial m^I(\theta^I)}{\partial \theta^{I'}}$ and $\frac{\partial}{\partial \theta^{I'}}\mathrm{vec}(\frac{\partial m^I(\theta^I)}{\partial \theta^{I'}})$ are $n_m \times n^I_\theta$ and $n_m \cdot n^I_\theta \times n^I_\theta$. The dimension of $F^I$ and $M^I$ are $n^I_\theta$ by $n^I_\theta$.

The theorem 3.1 in HMT states that the quasi-likelihood ratio test $T \cdot \mathrm{QLR}$ converges in distribution to Equation (18). The $n^I_\theta$ by $n^I_\theta$ matrix $V^I$ is defined as $V^I = V^I_1 - V^I_2 - V^I_3$ with $I = A, B$:

$$\begin{aligned} V^I_1 &= \frac{\partial m^I(\theta^I)}{\partial \theta^{I'}}(F^{I'})^{-1}\frac{\partial m^I(\theta^I)'}{\partial \theta^I} W \frac{\partial m^I(\theta^I)}{\partial \theta^{I'}}(F^I)^{-1}\frac{\partial m^I(\theta^I)'}{\partial \theta^I} \\ V^I_2 &= \frac{\partial m^I(\theta^I)}{\partial \theta^{I'}}((F^{I'})^{-1} + (F^I)^{-1})\frac{\partial m^I(\theta^I)'}{\partial \theta^I} \\ V^I_3 &= \frac{\partial m^I(\theta^I)}{\partial \theta^{I'}}(F^{I'})^{-1}(M^{I'} + M^I)(F^I)^{-1}\frac{\partial m^I(\theta^I)'}{\partial \theta^I} \end{aligned}$$

[12] We use the built-in procedures gradp and hessp in the GAUSS software package. The optimal step size for the gradient vector and the Hessian matrix is carefully adjusted because difference approximations might be imprecise as long as the first derivative is small. See Gill et al. (1981, Ch.4, pp. 127-133) for the choice of the finite-difference interval.

However, it is sometimes observed that the estimated $\hat{V}_B - \hat{V}_A$ is not a positive-definite matrix where some negative values are drawn in simulations. We should not discard the negative values of the test distribution when making statistical inference for the model comparison. The hypothesis test is assessed by critical values at the 1% and 5% confidence level $(Q_{99}, Q_{95})$ from the simulated asymptotic test distribution. When one model is nested within another, one rejects the null hypothesis at 5% level that two models are equivalent if $T \cdot \mathrm{QLR}(\widehat{\theta}_A, \widehat{\theta}_B) > Q_{95}$.

# F Simulated QLR distribution for model comparison

## F.1 Auto- and cross-covariances at lag 1: 15 moment conditions

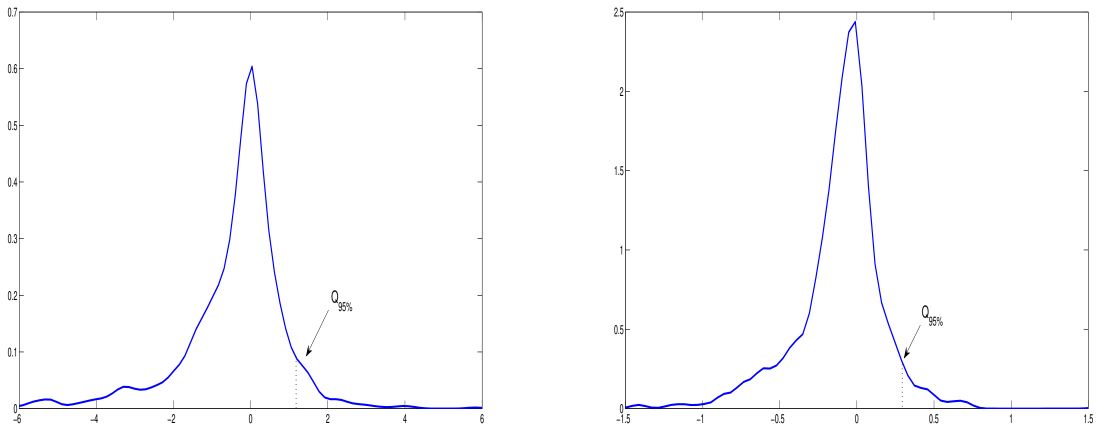


Figure 3: Test distribution for inflation persistence: GI (left) and GM (right)

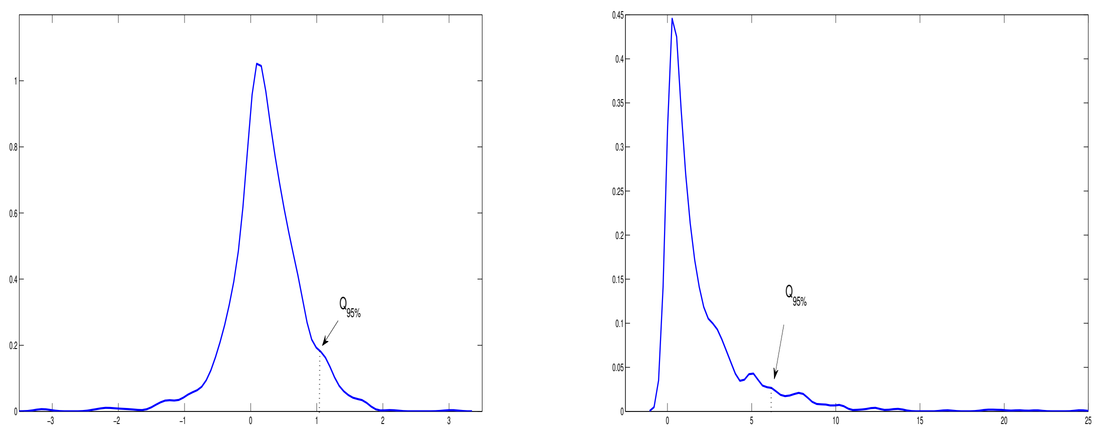


Figure 4: Test distribution for the output gap persistence: GI (left) and GM (right)

## F.2 Auto- and cross-covariances at lag 4: 42 moment conditions

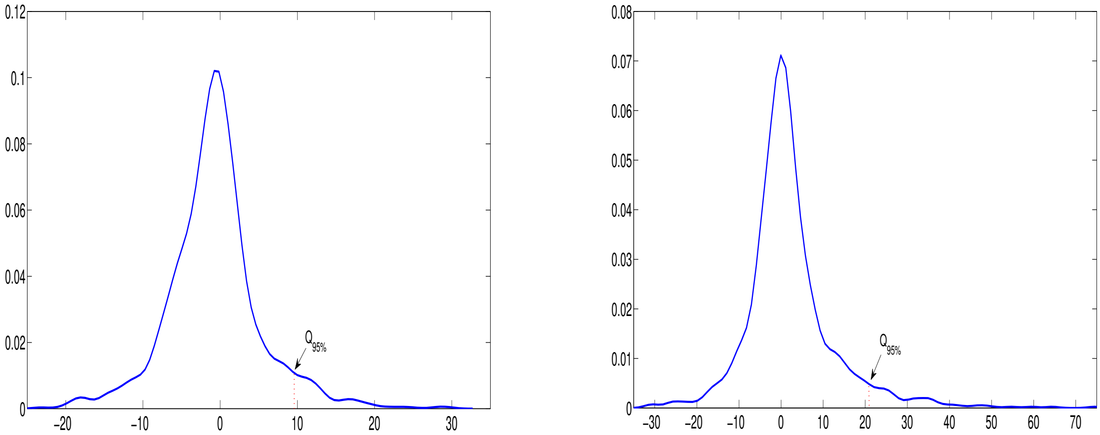


Figure 5: Test distribution for inflation persistence: GI (left) and GM (right)

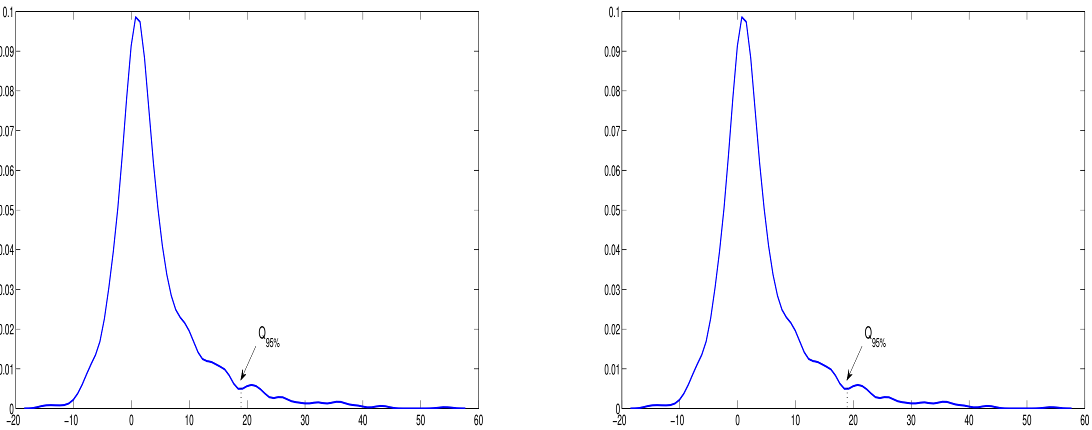


Figure 6: Test distribution for the output gap persistence: GI (left) and GM (right)

# G The Monte Carlo result of the misspecified case

Table G.2: Monte Carlo results on the MM and ML estimates of the misspecified model, ( ): RMSE, S.E : mean of standard errors

| | | ML | | | MM with 15 moments | | | MM with 42 moments | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\theta^0$ | T = 100 | T = 200 | T = 500 | T = 100 | T = 200 | T = 500 | T = 100 | T = 200 | T = 500 |
| $\kappa$ | 0.050 | 0.096 (0.186) | 0.089 (0.212) | 0.077 (0.031) | 0.176 (0.140) | 0.168 (0.125) | 0.163 (0.118) | 0.205 (0.175) | 0.191 (0.152) | 0.182 (0.136) |
| $\rho_\pi$ | 0.000 | 0.616 (0.621) | 0.618 (0.620) | 0.617 (0.618) | 0.632 (0.635) | 0.646 (0.647) | 0.653 (0.654) | 0.598 (0.604) | 0.614 (0.617) | 0.623 (0.624) |
| $\sigma_\pi$ | 0.675 | 0.491 (0.293) | 0.487 (0.330) | 0.474 (0.205) | 0.560 (0.151) | 0.543 (0.150) | 0.531 (0.654) | 0.661 (0.164) | 0.633 (0.127) | 0.612 (0.098) |
| $\chi$ | 1.000 | 0.921 (0.132) | 0.938 (0.100) | 0.955 (0.066) | 0.981 (0.053) | 0.994 (0.020) | 0.999 (0.015) | 0.970 (0.083) | 0.986 (0.047) | 0.997 (0.014) |
| $\tau$ | 0.090 | 0.085 (0.032) | 0.085 (0.024) | 0.085 (0.015) | 0.089 (0.029) | 0.086 (0.021) | 0.084 (0.014) | 0.088 (0.035) | 0.083 (0.024) | 0.080 (0.017) |
| $\sigma_x$ | 0.700 | 0.688 (0.064) | 0.691 (0.046) | 0.694 (0.026) | 0.637 (0.123) | 0.636 (0.103) | 0.636 (0.082) | 0.654 (0.132) | 0.644 (0.106) | 0.639 (0.083) |
| $\phi_\pi$ | 1.650 | 1.667 (0.182) | 1.657 (0.118) | 1.657 (0.075) | 1.691 (0.182) | 1.681 (0.117) | 1.679 (0.075) | 1.848 (0.291) | 1.783 (0.203) | 1.775 (0.156) |
| $\phi_x$ | 0.375 | 0.352 (0.127) | 0.352 (0.085) | 0.356 (0.054) | 0.227 (0.211) | 0.227 (0.203) | 0.226 (0.164) | 0.315 (0.282) | 0.238 (0.197) | 0.237 (0.166) |
| $\phi_r$ | 0.550 | 0.540 (0.049) | 0.541 (0.035) | 0.356 (0.054) | 0.488 (0.086) | 0.487 (0.077) | 0.489 (0.067) | 0.527 (0.070) | 0.524 (0.053) | 0.525 (0.038) |
| $\sigma_r$ | 0.750 | 0.738 (0.056) | 0.743 (0.039) | 0.748 (0.024) | 0.733 (0.101) | 0.744 (0.069) | 0.756 (0.043) | 0.597 (0.313) | 0.616 (0.244) | 0.649 (0.164) |
| $L(\theta)$ or $J(\theta)$ | | -398.38 | -805.68 | -2026.45 | 2.36 | 3.46 | 6.95 | 24.22 | 29.36 | 49.73 |

Note: The misspecified model does not include the parameter $\alpha$ in the NKPC. To save space, we also do not report the asymptotic standard errors for the parameter estimates, because these are not different from the case of the correctly specified case.