Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 143 (2008) 375–395
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Examining bias in estimators of linear rational expectations models under misspecification

Eric Jondeau[a], Hervé Le Bihan[b,*]

[a]*Faculty of Business and Economics, University of Lausanne and Swiss Finance Institute, Extranef 232, CH 1015 Lausanne, Switzerland*
[b]*Service de Recherche en Economie et Finance, Direction de la Recherche, Banque de France, 31 rue Croix des Petits Champs, 75049 Paris Cedex 01, France*

Available online 8 January 2008

Abstract

Most rational expectations models involve equations in which the dependent variable is a function of its lags and its expected future value. We investigate the asymptotic bias of generalized method of moment (GMM) and maximum likelihood (ML) estimators in such models under misspecification. We consider several misspecifications, and focus more specifically on the case of omitted dynamics in the dependent variable. In a stylized DGP, we derive analytically the asymptotic biases of these estimators. We establish that in many cases of interest the two estimators of the degree of forward-lookingness are asymptotically biased in opposite direction with respect to the true value of the parameter. We also propose a quasi-Hausman test of misspecification based on the difference between the GMM and ML estimators. Using Monte-Carlo simulations, we show that the ordering and direction of the estimators still hold in a more realistic New Keynesian macroeconomic model. In this set-up, misspecification is in general found to be more harmful to GMM than to ML estimators.

JEL: C13; C22; E31

Keywords: Rational expectations model; GMM estimator; ML estimator; Specification bias

1. Introduction

In this paper, we investigate the asymptotic bias of some well-known alternative estimators in rational expectations (RE) models under misspecification. We consider the so-called "hybrid" (or second-order) linear RE model

$$Y_t = \omega_f \mathrm{E}_t Y_{t+1} + \omega_b Y_{t-1} + \beta Z_t + \varepsilon_t, \tag{1}$$

where Y_t denotes the dependent variable, Z_t the forcing variable, ε_t the error term, and E_t the expectation conditional on the information available at date t. This framework has been widely used in theoretical as well as empirical work, because it provides a convenient framework for representing many macroeconomic

*Corresponding author.
E-mail addresses: eric.jondeau@unil.ch (E. Jondeau), herve.lebihan@banque-france.fr (H. Le Bihan).

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2007.11.004

behaviors. When expectations are unobserved, the estimation of (1) typically involves either the generalized method of moment (GMM) or a full-information maximum-likelihood estimation (MLE). With GMM, the expected variable is expressed as a function of an instrument set, without imposing more structure on the model. With MLE, the model is solved in terms of observed variables, taking into account the actual structure of the model.[1]

The general properties of GMM and ML estimators are rather well known for several kinds of misspecification. The asymptotic power properties of specification tests with GMM and MLE are investigated by Newey (1985a, b), respectively. White (1982) addresses the consequences of distributional misspecification on the ML estimator and examines how to detect such misspecification. Maasoumi and Phillips (1982) and Hall and Inoue (2003) derive the limiting distribution of instrumental variables (IV) and GMM estimators under the assumption of non-local misspecification, when the orthogonality conditions are not satisfied. To our knowledge, the comparison of GMM and ML estimators has not yet been performed in the context of misspecified RE models. As we will argue, the fact that in this context the endogenous regressor is the expected lead of the dependent variable has important consequences on the bias of the estimators. More specifically, we analyze the consequences of omitted dynamics on the asymptotic bias of both GMM and MLE in a very stylized set-up. In this context, these two approaches can be viewed as alternative procedures to perform the projection of the expectation terms onto a given information set. Within this framework, we derive the analytical expression of the asymptotic bias, and we show that the estimators of the forward-looking parameter ω_f are asymptotically biased *in opposite directions* with respect to the true value of the parameter.

The results presented in this paper help to rationalize some recent empirical evidence in the estimation of New Keynesian macroeconomic models. Key components of these models, such as the hybrid Phillips curve (e.g., Fuhrer, 1997; Galí and Gertler, 1999) and the hybrid aggregate demand equation (e.g., Fuhrer, 2000; Fuhrer and Rudebusch, 2004), are prominent illustrations of second-order RE models. It turns out that the GMM and MLE procedures produce a sizeable discrepancy in the estimation of the forward-looking parameter ω_f.[2] We show that the sign and the magnitude of the empirical discrepancy are actually consistent with our analytical results.

The structure of the paper is as follows. Section 2 presents a stylized hybrid RE model, describes the estimators, and investigates the finite-sample properties of GMM and ML estimators. Section 3 explores how misspecification does affect the asymptotic bias of these estimators. Several analytical results concerning the ranking of estimators are presented, and we propose a quasi-Hausman test for misspecification in RE models. In Section 4, we examine the properties of a multivariate New Keynesian model, inspired by Rudebusch (2002a, b), which allows to investigate omitted dynamics. Using Monte-Carlo simulations, we illustrate that the main result on the ordering of estimators carries on to this more realistic, analytically intractable, multivariate model. In that set-up, misspecification is also found to be more harmful to GMM estimators than to the ML estimator, in the sense that it yields larger biases on the forward-looking parameter ω_f. Section 5 provides concluding remarks.

2. A stylized hybrid RE model

In this section, we describe a simplified hybrid RE model, in which analytical results on the asymptotic bias of estimators under misspecification can be derived.

2.1. The model

The model includes a hybrid equation with one lag and one lead and the dynamics for the forcing variable as follows:

Model 1:

$$Y_t = \omega_f \mathrm{E}_t Y_{t+1} + (1-\omega_f) Y_{t-1} + \beta Z_t + \varepsilon_t, \tag{2}$$

$$Z_t = \rho Z_{t-1} + u_t, \tag{3}$$

[1] See for instance Pesaran (1987) for a review of both approaches.

[2] See, for instance, Jondeau and Le Bihan (2005) and Nason and Smith (2005) for the hybrid Phillips curve, and Fuhrer and Rudebusch (2004) for the aggregate demand equation.

where $0 \leqslant \omega_f \leqslant 1$ and $|\rho| < 1$. Error terms ε_t and u_t are contemporaneously and serially uncorrelated centered normal random variables, with $\sigma_\varepsilon^2 = \mathrm{E}[\varepsilon_t^2]$ and $\sigma_u^2 = \mathrm{E}[u_t^2]$. The sum of the forward-looking and backward-looking parameters is assumed to be equal to one, in accordance with many theoretical derivations in the New Keynesian macroeconomics (see, e.g., Christiano et al., 2005).[3] This assumption provides an identifying restriction required here given the AR(1) specification for the forcing variable.[4] In addition, it allows to obtain analytical expressions for the biases under misspecification. We assume for the moment that Z_t is strongly exogenous with respect to the parameters of the hybrid equation.

The properties of this model are derived from the characteristic polynomial of (2) given by $(1 - \omega_f L^{-1} - (1 - \omega_f)L) = 0$, with roots $\varphi_1 = (1 - \omega_f)/\omega_f$ and $\varphi_2 = 1$. According to the conditions for existence and uniqueness of solution to RE models, established by Blanchard and Kahn (1980), two situations can be encountered. When $\omega_f \leqslant 0.5$ (i.e., $\varphi_1 \geqslant 1$), the solution is unique, but Y_t is a non-stationary process regardless of the dynamics of Z_t. When $\omega_f > 0.5$ (i.e., $\varphi_1 < 1$) and $|\rho| < 1$, existence of a stationary solution is guaranteed, but there are in fact infinitely many solutions characterized by sunspot shocks. A special solution obtains when sunspot shocks are disregarded. This so-called fundamental solution is given by

$$Y_t = \varphi_1 Y_{t-1} + \theta Z_t + \tilde{\varepsilon}_t, \tag{4}$$

with $\theta = \beta/(\omega_f(1 - \rho))$ and $\tilde{\varepsilon}_t = \varepsilon_t/\omega_f$. We also define $\tilde{\sigma}_\varepsilon^2 = \mathrm{E}[\tilde{\varepsilon}_t^2] = \sigma_\varepsilon^2/\omega_f^2$. In empirical work, the presence of sunspot shocks may be an issue. In our subsequent analysis, based on Monte-Carlo simulations, we will simulate a DGP that coincides with the fundamental solution. Finally, our maintained assumptions for Model 1 are $0.5 < \omega_f \leqslant 1$ (or, equivalently, $0 \leqslant \varphi_1 < 1$).[5]

2.2. GMM and ML estimators

The estimation of (2) is typically based on two alternative procedures, namely GMM and MLE, to cope with the correlation of the regressors with the error term $\varepsilon_t' = \varepsilon_t - \omega_f(Y_{t+1} - \mathrm{E}_t Y_{t+1})$. In this section, we describe very simplified versions of these estimators that sketch their actual working.

GMM relies on using a set of instruments W_t uncorrelated with the error term ε_t' but correlated with the endogenous regressor (Y_{t+1}) to form moment conditions:

$$\mathrm{E}[W_t \cdot (Y_t - \omega_f Y_{t+1} - (1 - \omega_f) Y_{t-1} - \beta Z_t)] = 0. \tag{5}$$

Since two parameters (ω_f and β) have to be estimated, at least two instruments are required to achieve identification. The optimal instrument set for (2) is given by the two regressors $\{Y_{t-1}, Z_t\}$ in the reduced form (4). In this context, additional instruments would be redundant (see Breuch et al., 1999). Since the model is just identified, the probability limits (Plims) of the estimators are directly obtained by solving the moment conditions (5). Although the GMM approach reduces to IV estimation in this framework, the multivariate extension in Section 4 resorts to the genuine GMM estimator, so that we stick to the label GMM throughout the paper.

The MLE relies on using (3) to solve (2) iteratively forward. Since shocks are assumed to be uncorrelated, the full-information MLE reduces to a two-step approach in this context. In the first step, parameters φ_1 and θ in (4) and ρ in (3) are estimated under the constraints $0 \leqslant \varphi_1 < 1$ and $|\rho| < 1$.[6] In the second step, estimators of ω_f and β are obtained using the relations $\omega_f = 1/(1 + \varphi_1)$ and $\beta = \theta\omega_f(1 - \rho)$. The ML estimator is obtained under the assumption that forecasts are fully model consistent. The crucial difference between GMM and MLE is that MLE imposes some constraints upon the way Y_{t+1} is projected onto the state variables,

[3]Parameters ω_f and β may be functions of "deeper" parameters that characterize the preferences and constraints of agents. The results in this paper can in many cases be cast in terms of these deep parameters. However, the mapping of deep parameters to ω_f and β is specific to each application, so that we focus on the latter for more generality.

[4]See Pesaran (1987) or Mavroeidis (2005) for a discussion of the identification issue in RE models.

[5]It may be argued that the stationarity assumption may itself induce another misspecification, that we do not explore in this paper, however. Notice that, for more general dynamics for the forcing variable, a stationary process can be attained for lower values of ω_f (see Section 4).

[6]When a constraint is binding, the other parameters are re-estimated accordingly. Typically, if φ_1 is freely estimated to be larger than 1, the constrained estimate is given by $\hat{\varphi}_1 = 1 - \varepsilon$ and $\hat{\theta}$ freely estimated, with ε a given small positive value.

depending on the dynamics of Z_t used to solve the model. In contrast, GMM does not impose any constraint of this type on the first-step regression.

In a correctly specified model, both GMM and MLE provide consistent estimators. Thus, two reasons may explain the discrepancy between GMM and ML estimators found in many empirical estimates: (1) differences in the finite-sample properties of the estimators in a correctly specified model; (2) misspecification, resulting in inconsistency of GMM as well as ML estimators. While the present paper primarily focuses on misspecification, recent work has suggested that the discrepancy found between GMM and ML estimators comes from finite-sample bias (Fuhrer and Rudebusch, 2004; Lindé, 2005) or from lack of GMM identification (Mavroeidis, 2005). Therefore, we first discuss and assess finite-sample and weak-instrument biases in our set-up.

2.3. Sources of discrepancy in a correctly specified model

There are essentially two sources of bias for the GMM estimator in a correctly specified model. The first source is weak-instrument relevance that occurs when the correlation between instrument and endogenous regressors is weak. The recent literature emphasizes that weak identification may be present even in large samples.[7] It has been shown (Bound et al., 1995) that under weak identification the IV (or GMM) estimator is biased toward the Plim of the (inconsistent) OLS estimator. The magnitude of the weak-instrument bias is governed by the concentration parameter, which measures the strength of the instruments.

The GMM estimator also suffers from finite-sample bias. The bias is in that case also in the direction of the OLS estimator. The magnitude of the bias depends on the sample size, the number of instruments, and the multiple correlation between the instruments and the endogenous regressors (Nelson and Startz, 1990; Buse, 1992; Bound et al., 1995). A related issue is instrument redundancy, which occurs when some instruments are not correlated with the endogenous regressor conditionally on the presence of other instruments. This issue has been analyzed by Breuch et al. (1999) and Hall et al. (2007).

In the context of RE models, an abundant literature has studied the finite-sample properties of the GMM estimator and shown that this estimator may be severely biased and widely dispersed in small samples (see, among others, Fuhrer et al., 1995; Hansen et al., 1996). The MLE may also suffer from finite-sample bias, but this bias is generally found to be negligible in RE models (see Fuhrer et al., 1995; Jondeau et al., 2004).

In this section, we briefly discuss weak-instrument relevance and investigate the finite-sample bias more in depth. Since both types of biases are known to point to the direction of the Plim of the OLS estimator, we begin with the following proposition, which gives the Plim of the (inconsistent) OLS estimator obtained from the regression of Y_t onto $\{Y_{t+1}, Y_{t-1}, Z_t\}$ where the expected lead is replaced by the actual lead. It also gives the expression obtained in our set-up for the concentration parameter, which governs the magnitude of the GMM estimator bias under weak instruments.[8]

Proposition 1 (*Plim of the OLS estimator*). *Let us assume that the DGP is given by Model* 1. *The* (*inconsistent*) *OLS estimator obtained from* (2) *has the following Plim*:

$$\omega_{\mathrm{OLS}} = \frac{1}{2}\left(\frac{\tilde{\sigma}_\varepsilon^2 + \Lambda^2(1-\varphi_1)}{\tilde{\sigma}_\varepsilon^2 + \frac{1}{2}\Lambda^2[(1-\varphi_1^2) + (1-\varphi_1\rho)^2]}\right), \tag{6}$$

$$\beta_{\mathrm{OLS}} = \beta(1+\varphi_1)\frac{2-\varphi_1-\varphi_1\rho}{2(1-\varphi_1\rho)}\left(\frac{\tilde{\sigma}_\varepsilon^2 + \Lambda^2(1-\varphi_1\rho)}{\tilde{\sigma}_\varepsilon^2 + \frac{1}{2}\Lambda^2[(1-\varphi_1^2) + (1-\varphi_1\rho)^2]}\right), \tag{7}$$

[7] See Bound et al. (1995), Staiger and Stock (1997), Stock and Wright (2000), Stock et al. (2002). Staiger and Stock (1997) propose an asymptotic framework for IV, while Stock and Wright (2000) extend the approach to GMM. Methods robust to weak instruments have been proposed by Dufour (1997), Stock and Wright (2000), or Kleibergen (2002) and applied to the New Phillips curve by Nason and Smith (2005) or Dufour et al. (2006).

[8] Staiger and Stock (1997) show that $(1+\lambda_T^2/K)^{-1}$ approximates the magnitude of the finite-sample bias of IV relative to OLS estimator, where K is the number of instruments, and λ_T the concentration parameter.

with $\Lambda = \theta\sigma_u/(1-\varphi_1\rho)$. For the instrument set $\{Y_{t-1}, Z_t\}$, the concentration parameter λ_T is asymptotically given by

$$\frac{\lambda_T^2}{T} = \frac{(1-\varphi_1^2)(\tilde{\sigma}_\varepsilon^2 + \Lambda^2)}{(1+\varphi_1^2)\tilde{\sigma}_\varepsilon^2 + (1-\varphi_1\rho)^2\Lambda^2}.$$

Proof. See the Appendix.

The expression for the concentration parameter points out that weak identification is likely to occur when φ_1 is close to 1, or equivalently ω_f close to 0.5. In this case, which corresponds to a non-stationary dynamics for Y_t, the Plim of the OLS estimator is always smaller than 0.5. Given that the true parameter ω_f is above 0.5, and that the GMM estimator is biased in the direction of the Plim of the OLS estimator, the GMM estimator of ω_f is here biased downward. When $\varphi_1 < 1$, weak identification is asymptotically precluded in our framework.[9] However, even in this case, the finite-sample bias of the GMM estimator points towards the Plim of the OLS estimator. Contemplating (6), we observe that $0 \leqslant \omega_{\mathrm{OLS}} \leqslant 0.5$, so that the GMM estimator of ω_f is still biased downward.

To investigate the finite-sample bias more closely, we perform Monte-Carlo simulations of DGP (2)–(3). Concerning the GMM approach, we consider two instrument sets. The first one is the optimal instrument set, $W_t = \{Y_{t-1}, Z_t\}$, illustrating genuine finite-sample bias. The second set includes L additional (redundant) lags of W_t, illustrating instrument redundancy. It is worth emphasizing that it is a common practice in the empirical literature to include many instruments, in particular several lags. In this over-identified case, the two-stage least-square (TSLS) approach yields a consistent estimator, but designing an optimal weighting matrix requires acknowledging the MA(1) structure of the GMM error term, which results from one-period ahead expectations. We investigate two alternative approaches: (1) the well-known Newey and West (1987) weighting matrix with 12 lags (as typically used in Galí and Gertler, 1999), and (2) the efficient West (1997) matrix that assumes the error term to be an MA(1) process.

Table 1 reports the results of the simulation experiment. First, the ML estimator of ω_f is essentially unbiased, with a low dispersion. The estimator of β is unbiased for small values of ρ, while there is a slight positive bias when ρ is large.[10] Second, when the optimal instrument set $\{Y_{t-1}, Z_t\}$ is chosen ($L = 0$), the GMM estimator of ω_f is not significantly biased, although with a larger dispersion than that of the ML estimator. The magnitude of the bias increases with the extent of instrument redundancy (as in Hall et al., 2007). When the instrument set includes $L = 7$ redundant lags, the bias in the GMM estimator of ω_f is clearly negative. The higher the true value of ω_f, the larger the magnitude of the bias. The median estimate always points towards the Plim of the OLS estimator, which is below the true value of the parameter, and the magnitude of the bias increases with the value of the true parameter ω_f.

The GMM estimator biases under redundant instruments confirm and rationalize previous simulation evidence reported for instance by Fuhrer and Rudebusch (2004), Lindé (2005), or Mavroeidis (2005). It is worth noting, however, that the finite-sample bias in ω_f is not likely to account for the empirical conflict between GMM and ML estimates of hybrid models. Indeed most GMM estimates reported in the empirical literature point to a large weight on forward-looking expectations relative to lagged terms, suggesting a bias, if any, towards large values of ω_f. In contrast, in our set-up the finite-sample bias in the GMM estimator is towards the lower bound 0.5.

3. Asymptotic biases in misspecified models

This section investigates the consequences of misspecification on the asymptotic bias of GMM and ML estimators, with a focus on omitted dynamics in the dependent variable. Omitted dynamics has been pointed

[9] In a model where ω_f and ω_b are not assumed to sum to one, Mavroeidis (2005) identifies additional sources of under-identification that are related in particular to the dynamics of the forcing variable.

[10] For $\rho = 0.9$, the median estimate of β exceeds 0.1, whatever the true parameter ω_f. This bias is related to the downward bias of OLS in an autoregressive model. Since $\beta = \theta\omega_f(1-\rho)$, the negative bias in ρ translates in a positive bias in β.

Table 1
Finite-sample properties of estimators in the model with a single lag

Structural parameters		Statistic	GMM ($L=0$) (optimal instr. set)		GMM ($L=7$) (NW-12 lags)		GMM ($L=7$) (West-MA(1))		ML		Plim OLS	
ω_f	ρ		ω_f	β	ω_f	β	ω_f	β	ω_f	β	ω_f	β
0.55	0.0	Median	0.550	0.099	0.523	0.100	0.526	0.102	0.551	0.099	0.492	0.109
		MAD	0.049	0.063	0.034	0.063	0.047	0.071	0.014	0.045		
0.55	0.9	Median	0.548	0.108	0.534	0.122	0.534	0.124	0.550	0.106	0.459	0.198
		MAD	0.022	0.045	0.021	0.052	0.023	0.055	0.002	0.038		
0.75	0.0	Median	0.749	0.103	0.690	0.102	0.691	0.106	0.750	0.101	0.498	0.111
		MAD	0.074	0.088	0.059	0.084	0.069	0.094	0.043	0.062		
0.75	0.9	Median	0.746	0.106	0.699	0.109	0.700	0.111	0.751	0.106	0.488	0.132
		MAD	0.054	0.039	0.047	0.041	0.050	0.044	0.020	0.038		
0.95	0.0	Median	0.946	0.100	0.852	0.098	0.855	0.101	0.951	0.097	0.500	0.102
		MAD	0.106	0.114	0.089	0.105	0.099	0.122	0.072	0.080		
0.95	0.9	Median	0.947	0.105	0.855	0.103	0.859	0.107	0.954	0.105	0.499	0.105
		MAD	0.088	0.040	0.075	0.040	0.082	0.046	0.049	0.037		

This table reports summary statistics on the finite-sample distribution of the GMM and ML estimators of the model with a single lag. Parameter sets are $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0; 0.9\}$, $\beta = 0.1$, and $\sigma_\varepsilon = \sigma_u = 1$. This table reports the median and the median of absolute deviations from the median (MAD) of the empirical distribution obtained using Monte-Carlo simulations (multiplied by 1.4826 for comparability with a standard deviation). The experiment is performed as follows. For each parameter set, we simulate 5,000 samples of size $T = 150$. For each simulated sample, a sequence of $T + 100$ random innovations are drawn from the Gaussian distribution $\mathrm{N}(0, \Sigma)$ with no serial correlation ($\Sigma = \mathrm{diag}(\sigma_\varepsilon^2, \sigma_u^2)$), and the first 100 entries are discarded to reduce the effect of initial conditions. It also displays the Plim of the OLS estimator, towards which the GMM estimator is biased, under weak-instrument relevance or instrument redundancy. For the first GMM estimator, the instrument set only includes the optimal instruments $\{Y_{t-1}, Z_t\}$. For the second and third GMM estimators, $L = 7$ redundant lags of these instruments are also included. The weighting matrix of the second estimator is based on the Newey and West (1987) procedure (with 12 lags), while the third estimator uses the West (1997) procedure, acknowledging the MA(1) structure of the error term.

out in several studies as a likely source of misspecification in hybrid RE models (see, e.g., Kozicki and Tinsley, 2002; Rudebusch, 2002a, b). Other forms of misspecification are discussed more briefly.

3.1. The case of omitted dynamics

To illustrate omitted dynamics, the DGP is now assumed to include two lags of the dependent variable: *Model* 2:

$$\begin{aligned} Y_t &= \omega_f \mathrm{E}_t Y_{t+1} + \omega_b^1 Y_{t-1} + (1 - \omega_f - \omega_b^1) Y_{t-2} + \beta Z_t + \varepsilon_t, \\ Z_t &= \rho Z_{t-1} + u_t, \end{aligned} \tag{8}$$

where the parameters pertaining to lags and lead of inflation sum to one (we define $\omega_b^2 = 1 - \omega_f - \omega_b^1$). The characteristic polynomial of (8) is $(1 - \omega_f L^{-1} - \omega_b^1 L - \omega_b^2 L^2) = 0$ with roots $\varphi_1 = (1 - \omega_f)/\omega_f$, $\varphi_2 = \omega_b^2/\omega_f$, and $\varphi_3 = 1$. Consequently, the reduced form of this model is

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \theta Z_t + \tilde{\varepsilon}_t, \tag{9}$$

where as before $\theta = \beta/(\omega_f(1 - \rho))$ and $\tilde{\varepsilon}_t = \varepsilon_t/\omega_f$ with $\tilde{\sigma}_\varepsilon^2 = \sigma_\varepsilon^2/\omega_f^2$. Stationarity conditions of (9) (with an exogenous stationary forcing variable) are known to be $1 - \varphi_1 - \varphi_2 > 0$, $1 + \varphi_1 - \varphi_2 > 0$, and $\varphi_2 > -1$. These conditions are equivalent to the following conditions for the structural parameters: $(2 - 3\omega_f - \omega_b^1) < 0$ and $-\omega_f < \omega_b^1 < 1$.

Fig. 1 displays the area where the reduced form (9) is stationary assuming $|\rho| < 1$. If we maintain the assumption that $\omega_f \leqslant 1$ (or, equivalently, $\varphi_1 \geqslant 0$) to be consistent with the theoretical derivation of the hybrid

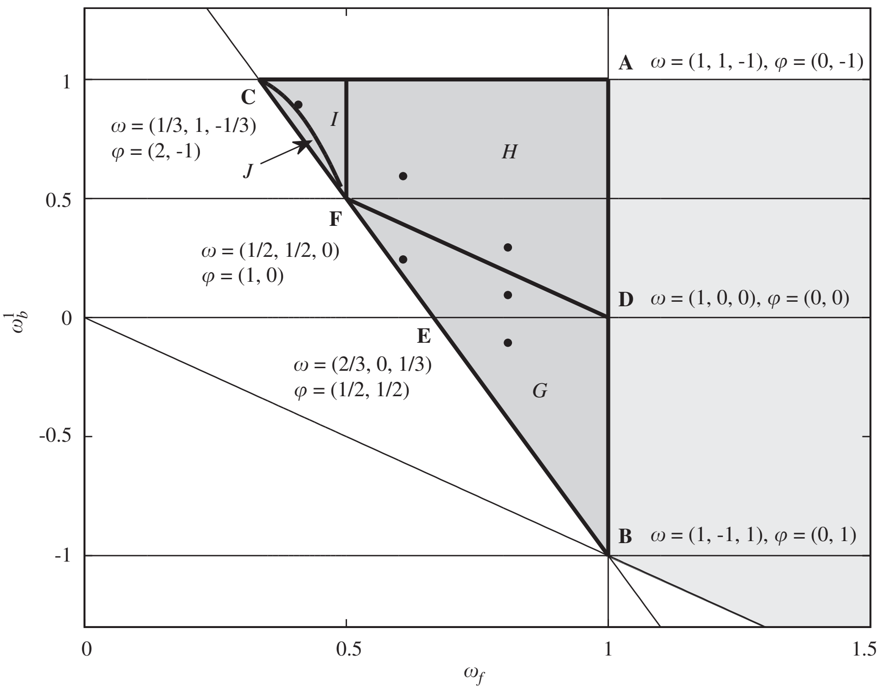

Fig. 1. This figure displays, in the plane $\{\omega_f, \omega_b^1\}$, the domain of validity of the hybrid model with two lags. The shaded area corresponds to the domain of stationarity of an AR(2) process. The area ABC corresponds to the additional constraint that $\omega_f \leqslant 1$. Area DEF corresponds to the domain where $0 \leqslant \omega_f, \omega_b^1, \omega_b^2 \leqslant 1$. The six points in the figure are the pairs $\{\omega_f, \omega_b^1\}$ selected for Table 2. Last, we denote $\omega = \{\omega_f, \omega_b^1, \omega_b^2\}$ and $\varphi = \{\varphi_1, \varphi_2\}$.

model, this yields the triangular area ABC. The segment BC corresponds to the non-stationarity case with $\varphi_1 + \varphi_2 = 1$. The segment AB is associated with $\omega_f = 1$. Along this segment, the degree of persistence $\varphi_1 + \varphi_2$ increases from-1 to 1. The segment CA corresponds to $\omega_b^1 = 1$, with a degree of persistence decreasing from 1 to -1.

The econometrician is assumed to erroneously select a single-lag specification, so that the estimated (misspecified) model is a one-lag hybrid model:

Model $2'$:

$$\begin{aligned} Y_t &= \alpha_f \mathrm{E}_t Y_{t+1} + (1 - \alpha_f) Y_{t-1} + bZ_t + v_t, \\ Z_t &= \rho Z_{t-1} + u_t. \end{aligned} \tag{10}$$

Estimators of parameter α_f are then used as estimators of the structural parameter ω_f. There is no misspecification in the limiting case where $\varphi_2 = 0$, i.e., $\omega_b^2 = 0$.

GMM and ML estimators are built as follows. As regards the selection of the GMM instrument set, we consider two alternative cases that may be reflective of actual practice. The first estimator (GMM1) is based on the instrument set $\{Y_{t-1}, Z_t\}$ that corresponds to the optimal instrument set in the (misspecified) DGP perceived by the econometrician. The second estimator (GMM2) resorts to the wider instrument set $\{Y_{t-1}, Y_{t-2}, Z_t\}$, including the omitted variable Y_{t-2}. This is the optimal instrument set in the correctly specified DGP. In both cases, however, the instruments are not valid for estimating (10) under DGP (8), since they are correlated with the error term v_t. The ML estimator is obtained by estimating by OLS the reduced form of the postulated Model $2'$, that is

$$Y_t = \phi Y_{t-1} + \mu Z_t + \tilde{v}_t, \tag{11}$$

where $\phi = (1 - \alpha_f)/\alpha_f$, $\mu = b/(\alpha_f(1 - \psi))$, and $\tilde{v}_t = v_t/\alpha_f$. Details on the construction of these estimators are provided in the Appendix. The Plims (or "pseudo-true values") of each estimator are summarized in Proposition 2.

Proposition 2 (*Plim of estimators with omitted dynamic*). *Assume that the DGP is given by Model 2 and that the misspecified Model 2′ is estimated. Then, the three estimators have the following Plims*:

- *GMM estimator with instrument set* $\{Y_{t-1}, Z_t\}$ (*GMM*1):

$$\alpha_{\mathrm{GMM1}} = \left(\frac{1}{1+\varphi_1-\varphi_2}\right)\left(\frac{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2[1+\varphi_2\rho(-1+\varphi_1-\varphi_2+\varphi_2\rho)]}{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2[1+\varphi_2\rho(\varphi_1+\varphi_2\rho)]}\right),$$

$$b_{\mathrm{GMM1}} = \beta\left(\frac{(1+\varphi_1)(1-\varphi_2-\rho(\varphi_1+\varphi_2+\varphi_2\rho))}{(1-\varphi_1\rho-\varphi_2\rho^2)(1+\varphi_1-\varphi_2)}\right) \times\left(\frac{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho(\varphi_1+\varphi_2\rho)\dfrac{2-\varphi_1\rho-\varphi_2\rho^2+\rho}{1-\varphi_2-\rho(\varphi_1+\varphi_2+\varphi_2\rho)}\right]}{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2[1+\varphi_2\rho(\varphi_1+\varphi_2\rho)]}\right).$$

- *GMM estimator with the instrument set* $\{Y_{t-1}, Y_{t-2}, Z_t\}$ (*GMM*2):

$$\alpha_{\mathrm{GMM2}} = \left(\frac{1-\varphi_1-\varphi_1\varphi_2}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right)\left(\frac{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\dfrac{-1+\varphi_1-\varphi_2+\varphi_2\rho}{1-\varphi_1-\varphi_1\varphi_2}\right]}{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\dfrac{-2\varphi_1\varphi_2+\varphi_2\rho-\varphi_2^2\rho}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right]}\right),$$

$$b_{\mathrm{GMM2}} = \beta\left(\frac{(1+\varphi_1)(1-\varphi_1-\varphi_2-\rho(1-\varphi_1-\varphi_1\varphi_2)(\varphi_1+\varphi_2+\varphi_2\rho))}{(1-\varphi_1\rho-\varphi_2\rho^2)(1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2)}\right) \times\left(\frac{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\dfrac{\varphi_1(1+\rho)(1-\varphi_1-\varphi_2)+\varphi_2\rho(2+\rho-2\varphi_1\rho-\varphi_2\rho)}{1-\varphi_1-\varphi_2-\rho(1-\varphi_1-\varphi_1\varphi_2)(\varphi_1+\varphi_2+\varphi_2\rho)}\right]}{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\dfrac{-2\varphi_1\varphi_2+\varphi_2\rho-\varphi_2^2\rho}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right]}\right).$$

- *ML estimator*:

$$\alpha_{\mathrm{ML}} = \left(\frac{1-\varphi_2}{1+\varphi_1-\varphi_2}\right)\left(\frac{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\dfrac{2\varphi_1+\varphi_2\rho(1-\varphi_2)}{1-\varphi_2}\right]}{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2[1+\varphi_2\rho(1+\varphi_1+(1+\rho)\varphi_2)]}\right),$$

$$b_{\mathrm{ML}} = \beta\left(\frac{(1+\varphi_1)(1-\varphi_2-\rho\varphi_1)}{(1-\varphi_1\rho-\varphi_2\rho^2)(1+\varphi_1-\varphi_2)}\right)\left(\frac{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\dfrac{2\varphi_1-\varphi_1^2\rho+\varphi_2\rho-\varphi_1\varphi_2\rho^2-\rho}{1-\varphi_2-\rho\varphi_1}\right]}{\tilde{\sigma}_\varepsilon^2+\tilde{\Lambda}^2[1+\varphi_2\rho(1+\varphi_1+(1+\rho)\varphi_2)]}\right),$$

with $\tilde{\Lambda} = \theta\sigma_u/(1-\varphi_1\rho-\varphi_2\rho^2)$.

Proof. See the Appendix.

Parameter $\tilde{\Lambda}$ is directly related to the covariance between Y_t and Z_t, since $\tilde{\Lambda}^2 = (1-\rho^2)\mathrm{E}[Z_tY_t]^2/V[Z_t]$. When $\varphi_2 = 0$, all estimators are asymptotically unbiased, with $\alpha_{\mathrm{GMM1}} = \alpha_{\mathrm{GMM2}} = \alpha_{\mathrm{ML}} = 1/(1+\varphi_1) = \omega_f$. Restricting to the case of an iid forcing variable ($\rho = 0$), we then obtain the following corollary.[11]

[11]The same results arise when $\tilde{\sigma}_\varepsilon^2$ is large relative to $\tilde{\Lambda}^2$.

Corollary 1. *In the case $\rho = 0$, Plims of estimators of α_f and b are given by*

$$\alpha_{\text{GMM1}} = \frac{1}{1+\varphi_1-\varphi_2}, \quad b_{\text{GMM1}} = \beta\left(\frac{(1+\varphi_1)(1-\varphi_2)}{1+\varphi_1-\varphi_2}\right),$$
$$\alpha_{\text{GMM2}} = \frac{1-\varphi_1-\varphi_1\varphi_2}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}, \quad b_{\text{GMM2}} = \beta\left(\frac{(1+\varphi_1)(1-\varphi_1-\varphi_2)}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right),$$
$$\alpha_{\text{ML}} = \frac{1-\varphi_2}{1+\varphi_1-\varphi_2}, \quad b_{\text{ML}} = \beta\left(\frac{(1+\varphi_1)(1-\varphi_2)}{1+\varphi_1-\varphi_2}\right).$$

Under stationarity ($-1<\varphi_1+\varphi_2<1$), the following inequalities hold:

$$1/2 \leqslant \alpha_{\text{ML}} \leqslant \omega_f \leqslant \alpha_{\text{GMM1}} \leqslant \alpha_{\text{GMM2}} \leqslant +\infty \quad \textit{if } \varphi_2 \geqslant 0 \ (\textit{area } \text{G}),$$
$$1/3 \leqslant \alpha_{\text{GMM1}} \leqslant \alpha_{\text{GMM2}} \leqslant \omega_f \leqslant \alpha_{\text{ML}} \leqslant 1 \quad \textit{if } \varphi_2 \leqslant 0 \textit{ and } \varphi_1 \leqslant 1 \ (\textit{area } \text{H}),$$
$$1/4 \leqslant \alpha_{\text{GMM1}} \leqslant \omega_f \leqslant \alpha_{\text{GMM2}} \leqslant \alpha_{\text{ML}} \leqslant 2/3 \quad \textit{if } \varphi_2 \in [-1;\bar{\varphi}] \textit{ and } \varphi_1 \geqslant 1 \ (\textit{area } \text{I}),$$
$$1/4 \leqslant \alpha_{\text{GMM1}} \leqslant \omega_f \leqslant \alpha_{\text{ML}} \leqslant \alpha_{\text{GMM2}} \leqslant 1/2 \quad \textit{if } \varphi_2 \in [\bar{\varphi};0] \textit{ and } \varphi_1 \geqslant 1 \ (\textit{area } \text{J}),$$

and

$$0 \leqslant b_{\text{GMM2}} \leqslant b_{\text{GMM1}} = b_{\text{ML}} \leqslant \beta \quad \textit{if } \varphi_2 \geqslant 0,$$
$$b_{\text{GMM2}} \leqslant \beta \leqslant b_{\text{GMM1}} = b_{\text{ML}} \quad \textit{if } \varphi_2 \leqslant 0,$$

where $\bar{\varphi} = (1-\varphi_1^2)/(1+\varphi_1^2-\varphi_1)$ is negative when $\varphi_1 \geqslant 1$.

We interpret the discrepancy of estimators obtained above as follows, assuming $\varphi_2 > 0$. As regards the GMM2 estimator, a crucial feature is that a relevant forcing variable is omitted from the estimated equation but included in the instrument set. As a result, the latter estimator suffers from an upward bias because $\hat{Y}_{t+1}$ partly captures the (positive) effect of the omitted variable Y_{t-2}. Concerning the GMM1 estimator, the omitted variable is not included in the instrument set. But given the persistence in the dependent variable, the first-step regression implies that $\widehat{Y}_{t+1}$ still captures a part of the omitted variable Y_{t-2}. Therefore, the bias in the GMM1 estimator of α_f is still positive. On the contrary, the variable Y_{t-2} is not included in the reduced-form equation (11) estimated for the ML estimator. As a consequence, Y_{t-1} partly captures the effect of Y_{t-2}. This creates a downward bias in the ML estimator of the forward-looking parameter, through the positive correlation of Y_{t-2} with Y_{t-1}.

The two largest areas (G and H) arguably reflect two well-known applications in the context of the New Keynesian macroeconomics.[12] In the New Keynesian Phillips curve, the additional lags are likely to have a positive cumulative effect so that reduced-form parameters are such that $\varphi_1 < 1$ and $\varphi_2 > 0$ (as in area G). In this context, the ML estimator of α_f is expected to have a downward bias, while the GMM estimator is expected to have an upward bias. In contrast, in the hybrid Euler equation for output gap, $\varphi_1 < 1$ and $\varphi_2 < 0$ (as in area H) are typically obtained as reduced-form parameters (e.g., Rudebusch, 2002a, b). In that case, the GMM and ML estimators point now in the opposite direction with respect to the true parameter ($\alpha_{\text{GMM}} \leqslant \omega_f \leqslant \alpha_{\text{ML}}$). Section 4 provides simulation evidence suggesting that the biases in the GMM and ML estimators of these two equations are actually in line with the analytical results.

Setting $\rho = 0$ may appear as an overly strong assumption, as compared with the serial correlation typically obtained for the forcing variable in many RE models. It is possible to show that the rank of the Plims of ML and GMM estimators is unaltered for most parameter sets. Although a formal proof of this result is not available, we checked it by computing the Plim of the various estimators using a grid for the structural parameters (using the expressions given in Proposition 2). For all values of ρ, as soon as $\omega_f \geqslant 0.5$, the ranking of GMM and ML estimators is $\alpha_{\text{ML}} \leqslant \alpha_{\text{GMM1}} \leqslant \alpha_{\text{GMM2}}$ when $\varphi_2 \geqslant 0$, and $\alpha_{\text{GMM1}} \leqslant \alpha_{\text{GMM2}} \leqslant \alpha_{\text{ML}}$ when $\varphi_2 \leqslant 0$.[13] Notice that, as confirmed by Fig. 1, the case $\omega_f \geqslant 0.5$ is the only case where we can consider small departure from the null hypothesis $\omega_b^2 = 0$ (corresponding to the segment DF). Indeed for $\omega_f < 0.5$, Y_t remains stationary only for strictly negative values of ω_b^2.

[12] Areas G–J are depicted in Fig. 1.

[13] When $\omega_f < 0.5$, the ranking depends on the analogous to parameter $\bar{\varphi}$, which in general depends on ρ in a very non-linear way.

3.2. Other sources of misspecification

Although the results presented in the previous section are specific to the DGP studied above, qualitatively similar results are obtained with other sources of misspecification. As an illustration, consider the case when a regressor is omitted in the hybrid equation. Such a case has been studied by Rudd and Whelan (2005) for the IV estimator in a backward-looking DGP. To generalize their framework, we consider the following hybrid DGP:

$$Y_t = \omega_f \mathrm{E}_t Y_{t+1} + (1-\omega_f) Y_{t-1} + \beta_1 Z_{1,t} + \beta_2 Z_{2,t} + \varepsilon_t, \tag{12}$$

$$Z_{1,t} = \rho_1 Z_{1,t-1} + u_{1,t}, \tag{13}$$

$$Z_{2,t} = \rho_2 Z_{2,t-1} + u_{2,t}, \tag{14}$$

with $\mathrm{E}(u_{1,t}u_{2,t}) = 0$, so that the two forcing variables are uncorrelated. The misspecified equation estimated by the econometrician is

$$Y_t = \alpha_f \mathrm{E}_t Y_{t+1} + (1-\alpha_f) Y_{t-1} + bZ_{1,t} + \varepsilon_t. \tag{15}$$

As shown in the Appendix, when $0 \leqslant \rho_1, \rho_2 \leqslant 1$, the following inequalities hold:

$$\alpha_{\mathrm{GMM1}} \leqslant \omega_f,$$
$$\alpha_{\mathrm{ML}} \leqslant \omega_f \leqslant \alpha_{\mathrm{GMM2}},$$

where α_{GMM1} and α_{GMM2} are the Plims of the GMM estimators with instrument sets $\{Y_{t-1}, Z_{1,t}\}$ and $\{Y_{t-1}, Z_{1,t}, Z_{2,t}\}$, respectively. As shown in a preliminary version of this paper, other misspecifications such as measurement error in the forcing variable would produce the same result that the GMM2 and the ML estimators are biased in opposite direction with respect to ω_f.[14]

In all these cases, a similar intuition applies: in the second-stage regression of the GMM2 estimation, the fitted lead of the dependent variable captures the effect of the omitted variable, so that parameter ω_f is over-estimated. On the opposite, with MLE, the effect of the omitted variable is captured by the lag of the dependent variable, resulting in parameter ω_f being under-estimated. Notice that in the omitted variable and measurement error cases, the GMM1 estimator of α_f always under-estimates ω_f. The reason is that the instrument set used in the first-step regression does not provide any useful information on the omitted variable that may be captured by the fitted lead of the dependent variable. By contrast, in the omitted dynamics case, the dependent variable, Y_{t-1}, does contain valuable information on the omitted lag Y_{t-2}.

3.3. A quasi-Hausman test for misspecification

The result that ML and GMM estimators generally diverge in opposite directions under misspecification suggests a Hausman-type test. The original Hausman (1978) specification test is based on the difference between an estimator consistent under the null and alternative hypotheses and an estimator efficient under the null hypothesis. In our context, under the null, the ML estimator is efficient while the GMM estimator is consistent. However, under the alternative of misspecification of the equation of interest, none of the estimators is consistent. We label this test a quasi-Hausman test.[15] As established in the previous sections, the Plims of the respective estimators differ under the alternative, so that the quasi-Hausman test is consistent for the types of misspecification considered here. It is likely that the test also has power against other misspecifications.

We define $\hat{\delta}_{\mathrm{ML}}$ and $\hat{\delta}_{\mathrm{GMM}}$ the estimators of $\delta = (\omega_f, \beta)'$ and V_{ML} and V_{GMM} their respective covariance matrices, estimated by $\hat{V}_{\mathrm{ML}}$ and $\hat{V}_{\mathrm{GMM}}$. Under the null that the DGP is correctly specified and assuming normality of errors, the estimator $\hat{\delta}_{\mathrm{ML}}$ is efficient. Then Hausman (1978)'s lemma applies.

[14] The role of measurement error in the new Phillips curve estimates is discussed by Lindé (2005).

[15] This terminology comes from Verbeek and Nijman (1992). Quasi-Hausman tests with the same feature have been proposed in other contexts, e.g., by White (1982) and Hahn and Hausman (2002).

Table 2
Plim of estimators and misspecification test in the case of omitted dynamics

Structural parameters					Panel A: plim of the estimator								Panel B: power of misspecification tests			
					GMM1		GMM2		Optimal GMM2		ML		With GMM2 (NW)		With GMM2 (W)	
ω_f	ω_b^1	ω_b^2	$\varphi_1+\varphi_2$	ρ	α_f	b	α_f	b	α_f	b	α_f	b	Hausman	J-stat	Hausman	J-stat
0.40	0.90	−0.30	0.75	0.00	0.308	0.135	0.526	0.053	0.530	0.048	0.538	0.135	0.446	1.000	0.178	1.000
0.40	0.90	−0.30	0.75	0.90	0.410	0.348	0.511	0.193	0.504	0.155	0.577	0.190	0.887	1.000	0.334	0.539
0.60	0.25	0.15	0.92	0.00	0.706	0.088	0.857	0.071	0.666	0.087	0.529	0.088	0.240	0.543	0.053	0.488
0.60	0.25	0.15	0.92	0.90	0.546	0.064	0.561	0.042	0.535	0.068	0.519	0.072	0.313	0.771	0.335	0.687
0.60	0.60	−0.20	0.34	0.00	0.500	0.111	0.536	0.107	0.523	0.118	0.667	0.111	0.878	0.896	0.869	0.889
0.60	0.60	−0.20	0.34	0.90	0.575	0.149	0.601	0.145	0.585	0.165	0.725	0.119	0.890	0.976	0.868	0.977
0.80	−0.10	0.30	0.62	0.00	1.143	0.089	1.217	0.087	0.961	0.090	0.714	0.089	0.478	0.877	0.199	0.855
0.80	−0.10	0.30	0.62	0.90	0.769	0.037	0.804	0.027	0.683	0.056	0.609	0.068	0.370	0.982	0.298	0.897
0.80	0.30	−0.10	0.12	0.00	0.727	0.102	0.730	0.102	0.676	0.103	0.818	0.102	0.387	0.245	0.345	0.207
0.80	0.30	−0.10	0.12	0.90	0.768	0.114	0.771	0.114	0.728	0.123	0.856	0.104	0.504	0.379	0.374	0.305
Bounds																
1.00	1.00	−1.00	−1.00	0.00	0.50		0.50		1.00		1.00					
1.00	1.00	−1.00	−1.00	1.00	0.50		0.50		1.00		1.00					
1.00	−1.00	1.00	1.00	0.00	Infinite		Infinite		1.00		1.00					
1.00	−1.00	1.00	1.00	1.00	0.50		Infinite		0.50		0.50					
0.33	1.00	−0.33	1.00	0.00	0.25		0.50		0.50		0.50					
0.33	1.00	−0.33	1.00	1.00	0.25		0.50		0.50		0.50					

This table reports the Plim of GMM and ML estimators (Panel A) and the power of misspecification tests (Panel B) in the case of omitted dynamics. Selected pairs of $\{\omega_f, \omega_b^1\}$ are displayed in Fig. 1. Other parameter sets are $\rho = \{0; 0.9\}$, $\beta = 0.1$ and $\sigma_\varepsilon = \sigma_u = 1$. The Plims are computed using Proposition 2. Bounds for the Plims of estimators of α_f and b are obtained for $\{\omega_f; \omega_b^1\} = \{A; B; C\}$ where A, B, and C are defined in Fig. 1, $\rho = \{0; 1\}$, and $\sigma_\varepsilon = 0$.

Proposition 3. *Under the null hypothesis of no misspecification, the statistic* $(\hat{\delta}_{\mathrm{GMM}} - \hat{\delta}_{\mathrm{ML}})' \hat{V}^{-1} (\hat{\delta}_{\mathrm{GMM}} - \hat{\delta}_{\mathrm{ML}})$ *is asymptotically distributed as a* χ^2 *with k degrees of freedom, where k is the dimension of* δ *and where* $\hat{V} = \hat{V}_{\mathrm{GMM}} - \hat{V}_{\mathrm{ML}}$.

In this proposition, the covariance matrix $\hat{V}$ is supposed to be positive definite. In some cases, it may be singular. Hausman and Taylor (1981) show that the same result applies provided one uses a generalized inverse of $\hat{V}$ in the estimation of the statistics.[16]

An additional feature is that in the case of misspecification affecting the ML auxiliary model (Eq. (3)), the GMM estimator is consistent under the alternative while the ML estimator is not. Rejection of the null hypothesis can be interpreted as a misspecification of either the equation of interest or the ML auxiliary equation.

3.4. Numerical evidence

To assess the quantitative importance of the estimator biases under misspecification, we present in Table 2 the Plims of GMM and ML estimators using formulae reported in Proposition 2. We select several pairs for $\{\omega_f, \omega_b^1\}$, corresponding to areas of interest in Fig. 1 (the selected points are displayed in the figure). The values correspond to a wide range of persistence $(\varphi_1 + \varphi_2)$ of the dependent variable. We also consider $\rho = \{0; 0.9\}$, $\beta = 0.1$ and $\sigma_\varepsilon = \sigma_u = 1$. (Results for other values are similar.) The values for ω_{b2} are chosen to be rather small, in order to cover moderate departures from model (2)–(3).

To obtain analytical results, Proposition 2 has abstracted from the issue of selecting the optimal weighting matrix. This is innocuous in the just-identified case (GMM1). However, in the over-identified case (GMM2), as put forward by Hall and Inoue (2003), different weighting matrices may yield different asymptotic biases under misspecification. In addition to the TSLS approach adopted in the previous sections, we thus also consider an estimation based on the efficient weighting matrix proposed by West (1997). In that case, the Plims are computed using a large-sample simulation.

Table 2 (Panel A) indicates that very large biases are likely to occur. Two important parameters affect the size of the bias, namely the size and the sign of the omitted lag (ω_b^2) and the persistence of the forcing variable (ρ). For all values of ρ, a negative parameter ω_b^2 tends to induce a negative bias in GMM estimators, but a positive bias in the ML estimator. In addition, the magnitude of this bias increases with ω_f for GMM, while it decreases with ω_f for MLE. With positive ω_b^2, in the case of no persistence $(\rho = 0)$, the ML bias is relatively small, while GMM estimators have a severe upward bias. For instance, when $\omega_f = 0.8$ and $\omega_b^1 = -0.1$ $(\omega_b^2 = 0.3)$, the pseudo-true values are $\alpha_{\mathrm{GMM1}} = 1.1$, $\alpha_{\mathrm{GMM2}} = 1.2$, and $\alpha_{\mathrm{ML}} = 0.7$. Even the optimal GMM estimator yields an estimate of $\alpha_{\mathrm{GMM2}} = 1$. In the high persistence case $(\rho = 0.9)$, the magnitude of the bias strongly reduces for GMM estimators and the bias turns out to be negative. On the contrary, the ML bias remains negative and increases in absolute value. It is worth emphasizing that in all cases considered the ranking of the GMM and ML estimators is not affected by introducing persistence in the forcing variable or varying the GMM weighting matrix (though the sign of the bias in α_f with respect to ω_f may be affected).

We turn now to the performances of the proposed quasi-Hausman test statistic relative to the widely used Hansen's (1982) J-statistic for misspecification (Panel B). We use a Monte-Carlo experiment with $T = 150$, a typical sample size with macroeconomic data. As before, we consider the GMM weighting matrices proposed by Newey and West (1987) and West (1997). Under the null hypothesis, the two test procedures perform rather well, their relative performance mainly depending on the weighting matrix used for the GMM estimation.[17] The table reports the power of the test statistics when the model (2)–(3) is actually misspecified, i.e., the percentage of the 5,000 replications in which the test statistic exceeds the relevant critical value of the χ^2

[16] In finite samples, it is also possible that some elements on the diagonal of $\hat{V}$ are estimated to be negative, because some parameters are more precisely estimated by GMM than by ML. In such instances, the Hausman test should not be interpreted, since we cannot conclude on the significance of the difference in point estimates.

[17] Unreported evidence shows that the tests based on the West (1997) weighting matrix are broadly correctly sized: for a nominal size of 5%, the actual rejection rate is typically between 3% and 10% for the J test and between 4% and 13% for the quasi-Hausman test. Tests based on the Newey and West (1987) weighting matrix provide higher rejection rates: between 5% and 14% for the J test and between 10% and 25% for the quasi-Hausman test. Detailed results are available upon request from the authors.

distribution for a nominal size of 5%. The two statistics display similar overall performances, and their relative performance depends on the parameter set. For some sets of parameters (typically, for $\omega_b^2 < 0$), the two statistics display a large power, suggesting that the statistics are able to discriminate these cases quite easily. In the case where ω_b^2 is positive, the two statistics have rather low power, the quasi-Hausman test performing the worst. The reason for the relative poor performance of the quasi-Hausman test lies in this case in the finite-sample bias already noticed. Indeed, for $\omega_b^2 > 0$, the GMM estimator is asymptotically biased upwards, but the finite-sample bias partially compensates this first effect, therefore reducing the gap between the two estimators.

4. A multivariate New Keynesian model

In the previous sections, we adopted a stylized model to provide analytical results under misspecification. In empirical work, however, the forcing variable is not likely to be strongly exogenous and may itself have a hybrid dynamics. Typical illustrations of such a multivariate hybrid model can be found in the New Keynesian macroeconomics (see Christiano et al., 2005). The general formulation of the multivariate hybrid RE model can be written in compact form as follows:

$$\Gamma_1 \mathbf{Y}_t = \Gamma_2 \mathrm{E}_t \mathbf{Y}_{t+1} + \Gamma_0 \mathbf{Y}_{t-1} + \boldsymbol{\eta}_t, \tag{16}$$

where $\mathbf{Y}_t$ is the vector of dependent variables and $\boldsymbol{\eta}_t$ the vector of shocks, while Γ_0, Γ_1 and Γ_2 are matrices that depend on structural parameters, denoted ξ. Shocks $\boldsymbol{\eta}_t$ are assumed to be serially uncorrelated, but they may be cross-correlated, so that their covariance matrix Σ is not necessarily diagonal. This model may accommodate dynamics including several lags or leads, provided $\mathbf{Y}_t$ is redefined accordingly. Using the undetermined coefficient approach, the structural form (16) has the following solution:

$$\mathbf{Y}_t = B_1 \mathbf{Y}_{t-1} + B_0 \boldsymbol{\eta}_t, \tag{17}$$

with $B_1 = (\Gamma_1 - \Gamma_2 B_1)^{-1}\Gamma_0$ and $B_0 = (\Gamma_1 - \Gamma_2 B_1)^{-1}$. Numerical procedures have been developed to compute the reduced form (17) of the model, such as the approach developed by Anderson and Moore (1985).[18] Once the reduced form is obtained, the ML estimator maximizes the concentrated log-likelihood function of the sample

$$\ln L(\xi) = -T[1 + \ln(2\pi)] - \frac{T}{2}\ln|\hat{\Sigma}| + \frac{T}{2}\ln|\hat{B}_0^{-1}|^2,$$

where $\hat{\Sigma} = T^{-1}\sum_{t=1}^{T} \hat{\boldsymbol{\eta}}_t(\xi)\hat{\boldsymbol{\eta}}_t(\xi)'$ is the estimated covariance matrix of residuals. Since no analytic solution can be obtained in this model, our analysis of the asymptotic bias is based on Monte-Carlo simulations.

4.1. The model

We focus on an empirical New Keynesian model inspired by Rudebusch (2002a, b) that is well suited to investigate explicitly the case of omitted dynamics.[19] This model combines a Phillips curve, an aggregate demand equation and a Taylor-rule type reaction function. It can be viewed as a multivariate extension of the model studied in the previous sections.

Model 3:

$$\pi_t = \omega_\pi \mathrm{E}_t \pi_{t+1} + (1-\omega_\pi)\sum_{j=1}^{4} \omega_{\pi b}^{j} \pi_{t-j} + \beta_y y_t + \varepsilon_{\pi,t}, \tag{18}$$

$$y_t = \omega_y \mathrm{E}_t y_{t+1} + (1-\omega_y)\sum_{k=1}^{2} \omega_{yb}^{k} y_{t-k} + \beta_r (i_t - \mathrm{E}_t \pi_{t+1}) + \varepsilon_{y,t}, \tag{19}$$

$$i_t = \rho i_{t-1} + (1-\rho)(\delta_\pi \pi_t + \delta_y y_t) + \varepsilon_{i,t}, \tag{20}$$

[18] An alternative estimation procedure has been recently put forward by Kurmann (2007) in a closely related context, following the approach originally developed by Sargent (1979).

[19] The reader may also refer to Kozicki and Tinsley (2002) for a review of interpretations of the hybrid Phillips curve with additional lags or to Fuhrer and Rudebusch (2004) for the hybrid aggregate demand equation with additional lags.

where π_t is the inflation rate, y_t the output gap and i_t the short-term nominal rate. $\varepsilon_{\pi,t}$, $\varepsilon_{y,t}$, and $\varepsilon_{i,t}$ denote error terms. We therefore define $\mathbf{Y}_t = (\pi_t, y_t, i_t)'$ and $\boldsymbol{\eta}_t = (\varepsilon_{\pi,t}, \varepsilon_{y,t}, \varepsilon_{i,t})'$.

The dynamics of inflation (18) is a hybrid Phillips curve with several own lags and the output gap as forcing variable. We impose the constraint $\sum_{j=1}^{4} \omega_{\pi b}^{j} = 1$ to be consistent with the natural rate hypothesis. Similarly, the aggregate demand equation for output gap (19) includes several own lags and the short-term nominal interest rate as forcing variable, with $\sum_{k=1}^{2} \omega_{yb}^{k} \leqslant 1$. Finally, (20) is a Taylor rule with interest-rate smoothing. It should be mentioned that this model includes the output gap as forcing variable in the Phillips curve, in place of the more theoretically grounded marginal cost, suggested by Galí and Gertler (1999). We adopt this specification because it provides a convenient multivariate representation for investigating the biases in the GMM and ML estimators. Another advantage of this model is that we can use the parameter estimates reported by Rudebusch (2002a, b) to calibrate our Monte-Carlo experiment. A multivariate model involving the marginal cost as forcing variable would raise be more problematic to calibrate, since no simple well-established model for the joint dynamics of inflation and marginal cost has been proposed to our knowledge.

Now, we explore the asymptotic bias of GMM and MLE in a misspecified model, in which only one lag is introduced in the dynamics of inflation and output gap. The assumed misspecified model is thus designed by setting $\omega_{\pi b}^{1} = \omega_{yb}^{1} = 1$ and $\omega_{\pi b}^{2} = \omega_{\pi b}^{3} = \omega_{\pi b}^{4} = \omega_{yb}^{2} = 0$ in Model 3:

Model 3′:

$$\pi_t = \alpha_\pi \mathrm{E}_t \pi_{t+1} + (1-\alpha_\pi)\pi_{t-1} + b_y y_t + v_{\pi,t}, \tag{21}$$

$$y_t = \alpha_y \mathrm{E}_t y_{t+1} + (1-\alpha_y) y_{t-1} + b_r(i_t - \mathrm{E}_t \pi_{t+1}) + v_{y,t}, \tag{22}$$

$$i_t = \rho i_{t-1} + (1-\rho)(\delta_\pi \pi_t + \delta_y y_t) + v_{i,t}. \tag{23}$$

4.2. Simulation results

The design of the Monte-Carlo simulation experiment relies on the model estimated by Rudebusch (2002a, b). For all parameters but ω_π and ω_y, we use the parameters reported by Rudebusch: $\omega_{\pi b}^{1} = 0.67$, $\omega_{\pi b}^{2} = -0.14$, $\omega_{\pi b}^{3} = 0.4$, $\omega_{\pi b}^{4} = 1 - \omega_{\pi b}^{1} - \omega_{\pi b}^{2} - \omega_{\pi b}^{3} = 0.07$, $\beta_y = 0.13$, $\omega_{yb}^{1} = 1.15$, $\omega_{yb}^{2} = -0.27$, $\beta_r = -0.09$, $\rho = 0.73$, $\delta_\pi = 1.53$, $\delta_y = 0.93$, and the standard deviations of innovations are $\sigma_{\varepsilon\pi} = 1.01$, $\sigma_{\varepsilon y} = 0.83$, and $\sigma_{\varepsilon i} = 0.36$. Then, we investigate several parameter sets for $\{\omega_\pi, \omega_y\}$ from 0.25 to 0.75.

Concerning the MLE, Model 3′ is estimated simultaneously, using the procedure described above. As regards GMM, the estimator follows the standards of the empirical literature and thus departs from the TSLS version used in previous section. Given the multivariate nature of the model and consistently with most previous work, valid instruments are assumed to be dated $t-1$ or earlier. The Phillips curve and the aggregate demand equation are estimated jointly in order to capture the possible correlation between moment conditions. The instrument set contains $\{\mathbf{Y}_{t-1}, \ldots, \mathbf{Y}_{t-4}\}$, in accordance with the number of lags in the reduced form of the correctly specified model (18)–(20).[20] The weighting matrix is computed using the West (1997) procedure. (Using the Newey–West weighting matrix yields similar results.)

In Table 3, we present two types of results. First, we report the Plim of the GMM and ML estimators, computed using a large sample of 100,000 observations. Second, we report the median and the MAD of the finite-sample distribution of the estimators, based on 5,000 samples of 150 observations. As regards GMM estimation, we obtain very large asymptotic biases when the forward-looking component in the Phillips curve is small to moderate. Whatever the value of the true parameter ω_π (from 0.25 to 0.75), the Plim of the GMM estimate ranges between 0.7 and 0.87. Even worse, unreported estimates indicate that, even when $\omega_\pi = 0.05$ (with $\omega_y = 0.5$), the Plim of the GMM estimate is still as high as 0.63. Thus GMM is not able to discriminate between a Phillips curves with a dominant backward-looking component and one with a dominant forward-looking component. In addition, in most cases, the output-gap parameter (b_y) is found to be negative although the true parameter is positive. The aggregate demand equation is much less prone to an asymptotic bias in the forward-looking parameter (α_y). The forward-looking component is not strongly affected, except in the case of

[20] We also investigated a GMM1-type estimator in which the lags of inflation and output gap beyond π_{t-1} and y_{t-1} were excluded from the instrument set. The results were not significantly altered, so that they are not reported to save space.

Table 3
Asymptotic and finite-sample properties of estimators in the hybrid multivariate model with omitted dynamics

Structural parameters		Statistic	GMM				ML				Power of misspecification tests	
ω_π	ω_y		α_π	b_y	α_y	b_r	α_π	b_y	α_y	b_r	Hausman	J-stat
0.50	0.50	Plim	0.751	−0.029	0.568	−0.069	0.543	−0.008	0.577	−0.052	0.745	0.972
		Median	0.690	0.001	0.559	−0.055	0.575	0.042	0.614	−0.044		
		MAD	0.048	0.041	0.026	0.028	0.030	0.021	0.023	0.013		
0.25	0.25	Plim	0.701	−0.008	0.484	−0.022	0.529	−0.002	0.328	−0.111	0.923	0.999
		Median	0.646	−0.002	0.490	−0.019	0.536	−0.003	0.293	−0.171		
		MAD	0.042	0.017	0.021	0.018	0.035	0.011	0.188	0.120		
0.75	0.75	Plim	0.869	−0.105	0.771	−0.085	0.764	0.068	0.786	−0.072	0.591	0.626
		Median	0.790	0.002	0.732	−0.076	0.768	0.081	0.791	−0.077		
		MAD	0.057	0.107	0.038	0.033	0.043	0.052	0.036	0.020		
0.25	0.75	Plim	0.747	−0.025	0.773	−0.057	0.618	−0.046	0.757	−0.063	0.872	0.928
		Median	0.667	−0.007	0.737	−0.055	0.606	−0.031	0.778	−0.063		
		MAD	0.050	0.027	0.038	0.027	0.048	0.040	0.029	0.010		
0.75	0.25	Plim	0.784	0.097	0.392	−0.096	0.745	0.091	0.397	−0.084	0.707	0.976
		Median	0.735	0.084	0.452	−0.053	0.742	0.092	0.391	−0.092		
		MAD	0.041	0.028	0.042	0.035	0.043	0.018	0.092	0.049		

This table reports estimates of the multivariate hybrid model, described in Eqs. (21)–(23), when it is estimated with a single lag. Parameters are those reported by Rudebusch (2002a, b): $\omega_{\pi b}^1 = 0.67$, $\omega_{\pi b}^2 = -0.14$, $\omega_{\pi b}^3 = 0.4$, $\omega_{\pi b}^4 = 1 - \omega_{\pi b}^1 - \omega_{\pi b}^2 - \omega_{\pi b}^3 = 0.07$, $\beta_y = 0.13$, $\omega_{yb}^1 = 1.15$, $\omega_{yb}^2 = -0.27$, $\beta_r = -0.09$, $\rho = 0.73$, $\delta_\pi = 1.53$, $\delta_y = 0.93$, $\sigma_{\varepsilon\pi} = 1.01$, $\sigma_{\varepsilon y} = 0.83$, and $\sigma_{\varepsilon i} = 0.36$. As in Rudebusch, we consider several values for ω_π and ω_y. The Plims of α_π, b_y, α_y and b_r are computed with a sample of 100,000 observations. Median and MAD correspond to the finite-sample distribution of α_π, b_y, α_y and b_r. They are obtained using Monte-Carlo simulation of 5,000 samples of size $T = 150$. Hausman and J-stat denote the power of the quasi-Hausman test and J test, respectively, i.e., the percentage of the 5,000 replications in which the test statistic exceeds the relevant critical value of the χ^2 distribution for a nominal size of 5%. The GMM estimation is performed using the West (1997) procedure for the weighting matrix (MA(1) error term). Eqs. (21)–(22) are jointly estimated. The instrument set contains four lags of the inflation rate, the output gap and the short-term interest rate. The ML estimates result from the joint estimation of Eqs. (21)–(23).

a very small value of the true parameter $\omega_y = 0.25$. In this case, the bias may be as high as 0.2, depending on the persistence in the Phillips curve.

Turning to the MLE, the forward-looking component of the Phillips curve is much less biased than the one obtained with GMM, so that the same ranking of Plims estimators as in Section 3 appears. Significant MLE biases are found for $\omega_\pi = 0.25$ only. When ω_π is equal to 0.5 or 0.75, the consequences of misspecification on α_π and α_y are modest, while estimates of b_y and b_r are correctly signed. Also, the finite-sample distribution of the ML estimator has a median very close to its Plim.

A noticeable feature of this experiment is that the discrepancies between GMM and ML estimators are close to those obtained on historical data.[21] For instance, the medians of the finite-sample distribution of the Phillips curve parameters reported in the table for $\{\omega_\pi; \omega_y\} = \{0.5; 0.5\}$ or $\{0.25; 0.25\}$ are close to the estimates reported by Jondeau and Le Bihan (2005) or Nason and Smith (2005). Indeed, the former obtain $\alpha_\pi = 0.61$ and $b_y = -0.025$ with GMM and $\alpha_\pi = 0.48$ and $b_y = 0.017$ with ML, while the latter report $\alpha_\pi = 0.68$ and $b_y = 0.008$ with GMM and $\alpha_\pi = 0.51$ and $b_y = -0.001$ with ML.

Finally, the table provides some evidence on the power of the quasi-Hausman and J statistics in the case of non-local misspecification. We observe that in all cases the two statistics perform rather well, since they are able to detect the assumed misspecification in at least 60% of the cases. In general, the J statistic has somewhat

[21] We have checked that in a correctly specified model (with $\omega_\pi = \omega_y = 0.5$), GMM and MLE are not likely to yield discrepancies between estimates of the forward-looking parameters as large as the ones obtained in a misspecified set-up. Typically, the probability of obtaining a gap of 0.12 (as reported in Table 3 for $\omega_\pi = \omega_y = 0.5$) is only 0.4%.

more power than the quasi-Hausman statistic. The table suggests that one reason for the lower power of the quasi-Hausman test is that the finite-sample bias on α_π systematically reduces, and sometimes offsets, the asymptotic bias. We observe the opposite for α_y, since the finite-sample bias exacerbates the asymptotic bias, but this effect is less sizeable than that of α_π. The quasi-Hausman test might have a better relative power under more local misspecifications or when restricting to a subset of parameters.

5. Conclusion

In this paper, we analyze the asymptotic biases in the GMM and ML estimators in misspecified second-order RE models. In the case of omitted dynamics, we show analytically that the GMM and ML biases generally point to opposite directions. The same contrast between the two estimators emerges when considering other plausible misspecifications or when a more elaborate multivariate model is considered.

Results in the present paper point to a critical source of discrepancy between estimators of a RE model, i.e., when a relevant regressor is omitted from the estimated equation but included in the GMM instrument set. Such an instance, rather likely when a large number of instruments is used, causes the lead of the dependent variable to capture the effect of the omitted variable, and the degree of forward-lookingness to be over-estimated. In the set-up analyzed here, misspecification of the equation of interest is typically found to be more harmful to the GMM estimator than to the MLE. This finding to some extent balances the well-known fact that in RE models, MLE may, unlike GMM, suffer from misspecification of the auxiliary model.

We also show through Monte-Carlo simulation of a multivariate New Keynesian model that GMM is more widely biased than MLE in a way that is likely to fill the gap between empirical estimates. Our results help to rationalize the empirical discrepancy reported between the large degree of forward-lookingness typically found when implementing GMM and the low degree of forward-lookingness obtained by MLE. Misspecification (and, in particular, omitted dynamics) typically induces biases in GMM and MLE that are consistent with the sign and the magnitude of parameter estimates reported in the empirical literature.

Acknowledgments

An earlier version of this paper has been presented at the Econometric Society European Meeting in August 2003 (Stockholm), at the North American Summer Meeting of the Econometric Society in June 2004 (Providence), and at seminars hold at the Banque de France, the ECB, the Université Paris 12 Val de Marne, GREQAM (Université de la Méditerranée), CORE (Catholic University of Louvain), and Oxford University. The authors are grateful to A. Zellner, an associate editor and three referees for very helpful comments. We also thank R. Davidson, P. Fève, A.M. Fuertes, C. Hurlin, A. Kurmann, F. Legendre, J. Matheron, S. Mavroeidis, and P. Sevestre for comments and discussions. This paper does not necessarily reflect the views of the Banque de France.

Appendix A

A.1. Model with a single lag

This section reports moments and cross-moments implied by Model 1. These expressions are used to compute the moment conditions involved in the definition of estimators.

A.1.1. Moments and cross-moments

Moments of Z_t are known to be $\mathrm{E}[Z_t^2] \equiv \sigma_Z^2 = \sigma_u^2/(1-\rho^2)$, and $\mathrm{E}[Z_t Z_{t-i}] = \rho^i \sigma_Z^2$. Cross-moments with Y_t are

$$\mathrm{E}[Z_t Y_t] = \frac{\theta\sigma_Z^2}{1-\varphi_1\rho} = \Gamma_0, \quad \mathrm{E}[Z_t Y_{t-1}] = \rho\Gamma_0 \quad \text{and} \quad \mathrm{E}[Y_t Z_{t-1}] = \varphi_1\Gamma_0 + \theta\rho\sigma_Z^2.$$

Moments involving Y_t only are

$$\begin{aligned}
\mathrm{E}[Y_t^2] &= \frac{\tilde{\sigma}_\varepsilon^2}{1-\varphi_1^2} + \frac{1+\varphi_1\rho}{1-\varphi_1^2}\theta\Gamma_0 = \Phi_0,\\
\mathrm{E}[Y_tY_{t-1}] &= \varphi_1\Phi_0 + \theta\rho\Gamma_0,\\
\mathrm{E}[Y_tY_{t-2}] &= \varphi_1^2\Phi_0 + \theta\rho(\varphi_1+\rho)\Gamma_0.
\end{aligned}$$

A.1.2. Proof of Proposition 1

Since (2) can be equivalently written as

$$Y_t - Y_{t-1} = \omega_f(\mathrm{E}_tY_{t+1} - Y_{t-1}) + \beta Z_t + \varepsilon_t, \tag{24}$$

the Plim of the OLS estimators of (2) is obtained by solving the following expressions:

$$\begin{aligned}
&\mathrm{E}[(Y_{t+1}-Y_{t-1})\cdot((Y_t-Y_{t-1}) - \omega_f(Y_{t+1}-Y_{t-1}) - \beta Z_t)] = 0,\\
&\mathrm{E}[Z_t\cdot((Y_t-Y_{t-1}) - \omega_f(Y_{t+1}-Y_{t-1}) - \beta Z_t)] = 0
\end{aligned}$$

yielding (6) and (7) in the text.

Rewriting (2) as in (24), we observe that the endogenous regressor is $(Y_{t+1}-Y_{t-1})$. Iterating (2) backward, we obtain the expression of the endogenous regressor as a function of the optimal instrument set (first-stage regression of the TSLS estimation)

$$Y_{t+1} - Y_{t-1} = (\varphi_1^2 - 1)Y_{t-1} + \theta_1(\varphi_1+\rho)Z_t + \tilde{\varepsilon}_{t+1} + \varphi_1\tilde{\varepsilon}_t + \theta u_t,$$

where the variance of the error term $v_t = \tilde{\varepsilon}_{t+1} + \varphi_1\tilde{\varepsilon}_t + \theta u_t$ is

$$\sigma_v^2 = (1+\varphi_1^2)\tilde{\sigma}_\varepsilon^2 + \theta^2\sigma_u^2.$$

Since there is one endogenous and one exogenous regressor in the equation, the concentration ratio is a scalar. Using the definition in Staiger and Stock (1997), it is then here given by

$$\frac{\lambda_T^2}{T} = \frac{\Omega(1-\varphi_1^2)^2}{\sigma_v^2},$$

where $\Omega = \mathrm{E}[Y_{t-1}^2] - \mathrm{E}[Z_tY_{t-1}]^2/\mathrm{E}[Z_t^2]$. Using the expressions above, we obtain

$$\frac{\lambda_T^2}{T} = \frac{(1-\varphi_1^2)(\tilde{\sigma}_\varepsilon^2 + \Lambda^2)}{(1+\varphi_1^2)\tilde{\sigma}_\varepsilon^2 + (1-\varphi_1\rho)^2\Lambda^2}.$$

A.2. Model with omitted dynamics

This section reports moments and cross-moments implied by Model 2 and describes how the Plims of the various estimators are computed.

A.2.1. Moments and cross-moments

Cross-moments between Z_t and Y_t are now given by

$$\begin{aligned}
\mathrm{E}[Z_tY_t] &= \frac{\theta\sigma_Z^2}{1-\varphi_1\rho-\varphi_2\rho^2} = \tilde{\Gamma}_0,\\
\mathrm{E}[Z_tY_{t-i}] &= \rho^i\tilde{\Gamma}_0, \quad \text{for all } i>0,\\
\mathrm{E}[Y_tZ_{t-1}] &= (\varphi_1+\varphi_2\rho)\tilde{\Gamma}_0 + \theta\rho\sigma_Z^2.
\end{aligned}$$

Moments involving Y_t only are

$$\mathrm{E}[Y_t^2] = \frac{(1-\varphi_2)\tilde{\sigma}_\varepsilon^2 + [\varphi_1\rho(1+\varphi_2) + (1-\varphi_2)(1+\varphi_2\rho^2)]\theta\tilde{\Gamma}_0}{(1+\varphi_2)(1-\varphi_1-\varphi_2)(1+\varphi_1-\varphi_2)} = \tilde{\Phi}_0,$$
$$\mathrm{E}[Y_t Y_{t-1}] = \frac{\varphi_1}{1-\varphi_2}\tilde{\Phi}_0 + \frac{\theta\rho}{1-\varphi_2}\tilde{\Gamma}_0,$$
$$\mathrm{E}[Y_t Y_{t-2}] = \left(\varphi_2 + \frac{\varphi_1^2}{1-\varphi_2}\right)\tilde{\Phi}_0 + \left(\frac{\varphi_1}{1-\varphi_2} + \rho\right)\theta\rho\tilde{\Gamma}_0.$$

A.2.2. Proof of Proposition 2

*Plim of GMM*1: Estimator GMM1 relies on the following moment conditions:

$$\mathrm{E}[Y_{t-1}\cdot(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f)Y_{t-1} - bZ_t)] = 0, \tag{25}$$
$$\mathrm{E}[Z_t\cdot(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f)Y_{t-1} - bZ_t)] = 0. \tag{26}$$

Since the model is just identified, the Plims of the estimators of α_f and b reported in Proposition 2 are obtained by solving the two moment conditions directly.

*Plim of GMM*2: Estimator GMM2 includes $\{Y_{t-1}, Y_{t-2}, Z_t\}$ with the omitted variable in the instrument set, leading to an over-identified parameter set. The estimator is built as a two-step estimator. First, Y_{t+1} is regressed on the instrument set to build the expectation of Y_{t+1} conditional on the information set, yielding $\hat{Y}_{t+1} = (\varphi_1^2+\varphi_2)Y_{t-1} + \varphi_1\varphi_2 Y_{t-2} + \theta(\varphi_1+\rho)Z_t$. Then, the Plims of estimators of α_f and b are obtained by solving the two following moment conditions:

$$\mathrm{E}[(\hat{Y}_{t+1} - Y_{t-1})\cdot(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f)Y_{t-1} - bZ_t)] = 0, \tag{27}$$
$$\mathrm{E}[Z_t\cdot(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f)Y_{t-1} - bZ_t)] = 0. \tag{28}$$

Plim of ML: The ML estimator is obtained by estimating the reduced form of the postulated Model 2′, that is (11). Parameters ϕ, μ, and ρ are estimated by OLS, and their Plims are denoted ϕ_{ML}, μ_{ML}, and ρ_{ML}. Then, the Plims of the ML estimators of α_f and b are given by the conditions $\alpha_{\mathrm{ML}} = 1/(1+\phi_{\mathrm{ML}})$ and $b_{\mathrm{ML}} = \mu_{\mathrm{ML}}\alpha_{\mathrm{ML}}(1-\rho_{\mathrm{ML}})$.

A.3. Model with omitted variable

This section describes how the Plims of the various estimators are computed when the DGP is given by (12)–(14), while the (misspecified) equation estimated by the econometrician is (15). The reduced form is

$$Y_t = \varphi_1 Y_{t-1} + \theta_1 Z_{1,t} + \theta_2 Z_{2,t} + \tilde{\varepsilon}_t,$$

where $\varphi_1 = (1-\omega_f)/\omega_f$, $\theta_1 = \beta_1/(\omega_f(1-\rho))$, $\theta_2 = \beta_2/(\omega_f(1-\rho))$, and $\tilde{\varepsilon}_t = \varepsilon_t/\omega_f$. We also define $\tilde{\sigma}_\varepsilon^2 = \mathrm{E}[\tilde{\varepsilon}_t^2] = \sigma_\varepsilon^2/\omega_f^2$.

A.3.1. Moments and cross-moments

Cross-moments between $Z_{i,t}$ and Y_t are given by, for $i = 1, 2$:

$$\mathrm{E}[Z_{i,t}Y_t] = \frac{\theta_i\sigma_{Zi}^2}{1-\varphi_1\rho_i} = \breve{\Gamma}_{0i},$$
$$\mathrm{E}[Z_{i,t}Y_{t-1}] = \rho_i\frac{\theta_i\sigma_{Zi}^2}{1-\varphi_1\rho_i} = \rho_i\breve{\Gamma}_{0i},$$
$$\mathrm{E}[Y_t Z_{i,t-1}] = \rho_i\breve{\Gamma}_{0i} + \theta_i\rho_i\sigma_{Zi}^2,$$

where $\sigma_{ui}^2 = \mathrm{E}[u_{i,t}^2]$ and $\sigma_{Zi}^2 = \sigma_{ui}^2/(1-\rho_i^2)$. Moments involving Y_t only are

$$\begin{aligned}
\mathrm{E}[Y_t^2] &= \frac{\tilde{\sigma}_\varepsilon^2}{(1-\varphi_1^2)} + \frac{(1+\varphi_1\rho_1)\theta_1^2\sigma_{Z1}^2}{(1-\varphi_1^2)(1-\varphi_1\rho_1)} + \frac{(1+\varphi_1\rho_2)\theta_2^2\sigma_{Z2}^2}{(1-\varphi_1^2)(1-\varphi_1\rho_2)} = \breve{\Phi}_0,\\
\mathrm{E}[Y_t Y_{t-1}] &= \varphi_1\breve{\Phi}_0 + \theta_1\rho_1\breve{\Gamma}_{01} + \theta_2\rho_2\breve{\Gamma}_{02},\\
\mathrm{E}[Y_t Y_{t-2}] &= \varphi_1^2\breve{\Phi}_0 + (\varphi_1+\rho_1)\theta_1\rho_1\breve{\Gamma}_{01} + (\varphi_1+\rho_2)\theta_2\rho_2\breve{\Gamma}_{02}.
\end{aligned}$$

A.3.2. Plims of the estimators under omitted variable

We first give the Plims of the estimators assuming model stationarity. Then, we derive some inequalities in the case of positive serial correlations and positive effect of the forcing variables.

*Plim of GMM*1: In the case of the just-identified GMM1 estimator with instrument set $\{Y_{t-1}, Z_{1,t}\}$, Plims of the estimators are obtained by solving the following moment conditions:

$$\mathrm{E}[Z_{1,t}\cdot(Y_t - \alpha Y_{t+1} - (1-\alpha)Y_{t-1} - bZ_{1,t})] = 0, \tag{29}$$
$$\mathrm{E}[Y_{t-1}\cdot(Y_t - \alpha Y_{t+1} - (1-\alpha)Y_{t-1} - bZ_{1,t})] = 0, \tag{30}$$

so that

$$\begin{aligned}
\alpha_{\mathrm{GMM1}} &= \left(\frac{1}{1+\varphi_1}\right)\frac{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{\theta_2^2\sigma_{u2}^2}{1-\varphi_1\rho_2}\dfrac{1}{1+\rho_2}}{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{\theta_2^2\sigma_{u2}^2}{1-\varphi_1\rho_2}\dfrac{1}{1-\rho_2^2}},\\
b_{\mathrm{GMM1}} &= \theta_1\frac{1-\rho_1}{1-\varphi_1\rho_1}(1-\alpha_{\mathrm{GMM1}}(1+\rho_1)\varphi_1),
\end{aligned}$$

where $\tilde{\Lambda}_1^2 = \theta_1^2\sigma_{u1}^2/(1-\varphi_1\rho_1)^2 = (1-\rho_1^2)\mathrm{E}[Z_{1,t}Y_t]^2/V[Z_{1,t}]$.

*Plim of GMM*2: In the case of the GMM2 estimator, based on the instrument set $\{Y_{t-1}, Z_{1,t}, Z_{2,t}\}$, we first compute the fitted value from the first-stage regression: $\hat{Y}_{t+1} = \varphi_1^2 Y_{t-1} + \theta_1(\varphi_1+\rho_1)Z_{1,t} + \theta_2(\varphi_2+\rho_2)Z_{2,t}$. Then, the Plims of the second-stage regression are the solution of the following moments conditions:

$$\mathrm{E}[Z_{1,t}\cdot(Y_t - \alpha\hat{Y}_{t+1} - (1-\alpha)Y_{t-1} - bZ_{1,t})] = 0, \tag{31}$$
$$\mathrm{E}[(\hat{Y}_{t+1} - Y_{t-1})\cdot(Y_t - \alpha\hat{Y}_{t+1} - (1-\alpha)Y_{t-1} - bZ_{1,t})] = 0, \tag{32}$$

so that

$$\begin{aligned}
\alpha_{\mathrm{GMM2}} &= \left(\frac{1}{1+\varphi_1}\right)\frac{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{\theta_2^2\sigma_{u2}^2}{1-\varphi_1\rho_2}\dfrac{1}{1-\varphi_1}}{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{\theta_2^2\sigma_{u2}^2}{1-\varphi_1\rho_2}\dfrac{1}{1-\varphi_1^2}},\\
b_{\mathrm{GMM2}} &= \theta_1\frac{1-\rho_1}{1-\varphi_1\rho_1}(1-\alpha_{\mathrm{GMM2}}(1+\rho_1)\varphi_1).
\end{aligned}$$

Plim of ML: The reduced form of the misspecified equation (15) is given by

$$Y_t = \phi_1 Y_{t-1} + \mu_1 Z_{1,t} + \tilde{\varepsilon}_t.$$

Estimators of ϕ_1 and μ_1 are therefore obtained from the two following moments conditions:

$$\mathrm{E}[Y_{t-1}\cdot(Y_t - \phi_1 Y_{t-1} + \mu_1 Z_{1,t})] = 0, \tag{33}$$
$$\mathrm{E}[Z_{1,t}\cdot(Y_t - \phi_1 Y_{t-1} + \mu_1 Z_{1,t})] = 0 \tag{34}$$

yielding the Plims

$$\phi_{\mathrm{ML}} = \frac{\varphi_1(\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2) + \dfrac{(\varphi_1 + \rho_2)}{(1-\varphi_1\rho_2)}\theta_2^2(1-\rho_2^2)\sigma_{u2}^2}{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{(1 + \varphi_1\rho_2)}{(1-\varphi_1\rho_2)}\theta_2^2(1-\rho_2^2)\sigma_{u2}^2},$$

$$\mu_{\mathrm{ML}} = \theta_1 + (\varphi_1 - \phi_{\mathrm{ML}})\frac{\rho_1\theta_1}{(1-\varphi_1\rho_1)}.$$

The Plims of the ML estimators of α_f and b are given by the conditions $\alpha_{\mathrm{ML}} = 1/(1+\phi_{\mathrm{ML}})$ and $b_{\mathrm{ML}} = \mu_{\mathrm{ML}}\alpha_{\mathrm{ML}}(1-\rho_{\mathrm{ML}})$:

$$\alpha_{\mathrm{ML}} = \left(\frac{1}{1+\varphi_1}\right)\frac{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{\theta_2^2\sigma_{u2}^2}{1-\varphi_1\rho_2}\dfrac{1+\varphi_1\rho_2}{1-\rho_2^2}}{\tilde{\sigma}_\varepsilon^2 + \tilde{\Lambda}_1^2 + \dfrac{\theta_2^2\sigma_{u2}^2}{1-\varphi_1\rho_2}\dfrac{1}{1-\rho_2}},$$

$$b_{\mathrm{ML}} = \theta_1\frac{1-\rho_1}{1-\varphi_1\rho_1}(\alpha_{\mathrm{ML}}(1+\rho_1) - \rho_1).$$

Ranking of the estimators: The ranking of the Plims of GMM and ML estimators is obtained in the typical case where the serial correlation of the forcing variables (ρ_1 and ρ_2) is positive.

Corollary 2. *When* $0 \leqslant \rho_1, \rho_2 \leqslant 1$, *the following inequalities hold*:

$$\alpha_{\mathrm{GMM1}} \leqslant \omega_f,$$
$$\alpha_{\mathrm{ML}} \leqslant \omega_f \leqslant \alpha_{\mathrm{GMM2}}.$$

In addition, *when* $\beta_2 > 0$, *the following inequalities hold*:

$$b_{\mathrm{GMM2}} \leqslant \beta_1,$$
$$b_{\mathrm{GMM1}} \leqslant \beta_1 \leqslant b_{\mathrm{ML}}.$$

As before, these inequalities are obtained by comparing terms between brackets in the numerator and denominator of each expression.

References

Anderson, G.A., Moore, G.R., 1985. A linear algebraic procedure for solving perfect foresight models. Economics Letters 17, 247–252.

Blanchard, O.J., Kahn, C.M., 1980. The solution of linear difference models under rational expectations. Econometrica 48, 1305–1311.

Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. Journal of the American Statistical Association 90, 443–450.

Breuch, T., Qian, H., Schmidt, P., Wyhowski, D., 1999. Redundancy of moment conditions. Journal of Econometrics 91, 89–111.

Buse, A., 1992. The bias of instrumental variable estimators. Econometrica 60, 173–180.

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

Dufour, J.M., 1997. Some impossibility theorems in econometrics with applications to structural and dynamic models. Econometrica 65, 1365–1388.

Dufour, J.M., Khalaf, L., Kichian, M., 2006. Inflation dynamics and the New Keynesian Phillips curve: an identification-robust econometric analysis. Journal of Economic Dynamics and Control 30, 1707–1727.

Fuhrer, J.C., 1997. The (un)importance of forward looking behavior in price specifications. Journal of Money, Credit, and Banking 29, 338–350.

Fuhrer, J.C., 2000. Habit formation in consumption and its implications for monetary-policy models. American Economic Review 90, 367–390.

Fuhrer, J.C., Rudebusch, G.D., 2004. Estimating the Euler equation for output. Journal of Monetary Economics 51, 1133–1153.

Fuhrer, J.C., Moore, G.R., Schuh, S.D., 1995. Estimating the linear quadratic inventory model. Maximum likelihood versus generalized method of moments. Journal of Monetary Economics 35, 115–157.

Galí, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. Journal of Monetary Economics 44, 195–222.

Hahn, J., Hausman, J.A., 2002. A new specification test for the validity of instrumental variables. Econometrica 70, 163–189.

Hall, A.R., Inoue, A., 2003. The large sample behavior of the generalized method of moments estimator in misspecified models. Journal of Econometrics 114, 361–394.

Hall, A.R., Inoue, A., Jana, K., Shin, C., 2007. Information in the generalized method of moments estimation and entropy based moment selection. Journal of Econometrics 138, 488–512.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimator. Econometrica 50, 1029–1054.

Hansen, L.P., Heaton, J., Yaron, A., 1996. Finite-sample properties of some alternative GMM estimators. Journal of Business and Economic Statistics 14, 262–280.

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.

Hausman, J.A., Taylor, W.E., 1981. A generalized specification test. Economics Letters 8, 239–245.

Jondeau, E., Le Bihan, H., 2005. Testing for a forward-looking Phillips curve. Additional evidence from European and US data. Economic Modelling 22, 521–550.

Jondeau, E., Le Bihan, H., Gallès, C., 2004. Assessing generalized method of moments estimates of the Federal Reserve reaction function. Journal of Business and Economic Statistics 22, 225–239.

Kleibergen, F., 2002. Pivotal statistics for testing structural parameters in instrumental variables regression. Econometrica 70, 1781–1804.

Kozicki, S., Tinsley, P.A., 2002. Alternative sources of the lag dynamics of inflation. Federal Reserve Bank of Kansas City Working Paper 02-12.

Kurmann, A., 2007. VAR based estimation of Euler equations, with an application to new Keynesian pricing. Journal of Economics Dynamics and Control 31, 767–796.

Lindé, J., 2005. Estimating New-Keynesian Phillips curves: a full information maximum likelihood approach. Journal of Monetary Economics 52, 1135–1142.

Maasoumi, E., Phillips, P.C.B., 1982. On the behavior of inconsistent instrumental variable estimators. Journal of Econometrics 19, 183–201.

Mavroeidis, S., 2005. Identification issues in forward-looking monetary models estimated by GMM, with an application to the New Phillips curve. Journal of Money, Credit, and Banking 37, 421–448.

Nason, J.M., Smith, G.W., 2005. Identifying the New Keynesian Phillips curve. Working Paper 2005-1, Federal Reserve Bank of Atlanta.

Nelson, C.R., Startz, R., 1990. Some further results on the exact small sample properties of the instrumental variables estimator. Econometrica 58, 967–976.

Newey, W.K., 1985a. Maximum likelihood specification testing and conditional moment tests. Econometrica 53, 1047–1070.

Newey, W.K., 1985b. Generalized method of moments specification testing. Journal of Econometrics 29, 229–256.

Newey, W.K., West, K.D., 1987. A simple, positive definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pesaran, H., 1987. The Limits to Rational Expectations. Basil Blackwell, Oxford.

Rudebusch, G.D., 2002a. Assessing nominal income rules for monetary policy with model and data uncertainty. Economic Journal 112, 402–432.

Rudebusch, G.D., 2002b. Term structure evidence on interest rate smoothing and monetary policy inertia. Journal of Monetary Economics 49, 1161–1187.

Rudd, J., Whelan, K., 2005. New tests of the New-Keynesian Phillips curve. Journal of Monetary Economics 52, 1167–1187.

Sargent, T.J., 1979. A note on maximum likelihood estimation of the rational expectations model of the term structure. Journal of Monetary Economics 5, 133–143.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.

Stock, J.H., Wright, J.H., 2000. GMM with weak identification. Econometrica 68, 1055–1096.

Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20, 518–529.

Verbeek, M., Nijman, T., 1992. Testing for selectivity bias in panel data models. International Economic Review 33, 681–704.

West, K.D., 1997. Another heteroskedasticity- and autocorrelation-consistent covariance matrix estimator. Journal of Econometrics 76, 171–191.

White, H., 1982. Maximum likelihood estimation of misspecified models. Econometrica 50, 1–25.