ML vs GMM Estimates of Hybrid Macroeconomic Models (With an Application to the "New Phillips Curve")

Eric Jondeau* Hervé Le Bihan†

February 2003
(Preliminary version)

Abstract

Many macroeconomic models involve hybrid equations, in which some variables are a function of both their lags and their expected future value. The hybrid "New Keynesian" Phillips Curve is a prominent example. In this context, hybrid models have produced conflicting empirical results: Studies which use ML estimation tend to find the forward-looking component to be small, while those using GMM have reported inflation dynamics to be predominantly forward-looking.

This paper provides a rationalization for this empirical conflict. Allowing for two alternative and straightforward kinds of mis-specifications (omitted dynamics and measurement error) in a hybrid model, we show that the ML estimator tends to underevaluate the degree of forward-lookingness, while the GMM estimator tends to overstate it. This result is analytically shown to hold in a simple DGP. Monte-Carlo experiments indicate that it remains valid in a wide range of plausible DGPs. Simulations also indicate that discrepancies of the size observed in the new Phillips curve context can more readily be accounted for by misspecification, than by the GMM small sample bias.

Keywords: Rational-expectation model, GMM estimator, ML estimator, Inflation, New Phillips curve.
JEL classification: E31.

*Banque de France, Centre de recherche, 41-1391 DGEI, 31 rue Croix des petits champs, 75049 Paris Cedex 01, e-mail: eric.jondeau@banque-france.fr, and ERUDITE, Université Paris 12 Val-de-Marne.

†Banque de France, Centre de recherche, e-mail: herve.lebihan@banque-france.fr.

1 Introduction

Many macroeconomic models involve estimating a "hybrid" equation of the form:

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

where Y_t denotes the dependent variable, Z_t the forcing variable, and ε_t the error term. Such a specification has been referred to as hybrid, because it nests a forward-looking specification often derived from a Euler condition as well as the backward-looking autoregressive distributed-lag specification.

A prominent example is the hybrid Phillips Curve, in which the inflation rate depends on its own lead and lag and on the real marginal cost. This model has been proposed by Fuhrer and Moore (1985b), Galí and Gertler (1999), and Christiano, Eichenbaum, and Evans (2001) to introduce some inflation persistence in the purely forward-looking model of Rotemberg (1982) and Calvo (1983). Other examples of specification (1) may be found in the literature on inventories (Fuhrer, Moore, and Schuh, 1995), on investment (Oliner, Rudebusch, and Sichel, 1996), on consumption (Fuhrer, 2000), or more recently on output gap (Fuhrer and Rudebusch, 2002). More generally, such a specification is typically the reduced form of rational-expectation models where the adjustment cost is defined on the change of the variable, proposed for instance by Tinsley (1993, 2000). Alternative derivations of the hybrid model, based on rule-of-thumb behavior, have also been emphasized by Amato and Laubach (2003).

Estimation of equation (1) under the rational-expectation assumption typically involves using either the Generalized-Method-of-Moment (GMM) or the Maximum-Likelihood (ML) approaches. GMM and ML are alternative procedures to cope with the unobserved forward-looking component of the hybrid model.[1] GMM expresses the expected variable $E_t Y_{t+1}$ as a function of an instrument set, without referring to the structure of the process for the forcing variable. ML produces model-consistent forecasts of Y_{t+1} in taking into account the structure of the equation for Z_t. While the two approaches are asymptotically equivalent, a recurrent finding is that empirical estimates of the hybrid model produce contrasting results, suggesting that the estimation method plays a role in the conflict. For instance, Fuhrer (1997), using the ML approach, has found the forward-looking component in US inflation to be essentially unimportant. On the other hand, Galí and Gertler (1999), using the GMM approach, have reported that the forward-looking component was dominant. Similarly, Fuhrer and Rudebusch (2002) have estimated a hybrid I-S curve using both estimation approaches, and have found the forward-looking parameter to be systematically larger when the estimation is performed with GMM than with ML.

The purpose of the present paper is to rationalize the discrepancy between empirical estimates obtained using the ML and GMM approaches, with a focus on the forward-looking parameter ω_f. This parameter is indeed crucial in many applications. In the new Phillips curve context, the value of this parameter has dramatic implications for policy purposes, since it directly affects the effectiveness of monetary policy (see Fuhrer 1997).

[1] The present paper focuses on comparing ML and GMM in this perspective: We do not put the emphasis on questions such as choosing the precise shape of the likelihood function or the optimal GMM weighting matrix. In particular, in the case we consider, GMM reduces to the two-stage least squares.

We first point that, unless a large number of irrelevant instruments is used, finite-sample biases are not likely to fill the gap between parameter estimates typically obtained in empirical applications. We then show that two natural mis-specifications (measurement error in the forcing variable and omitted dynamics) can produce large discrepancies. The results are established analytically in a stylized representative framework. Interestingly, in most cases, the probability limits of the GMM and ML estimators of the degree of forward-lookingness are biased *in opposite direction* from the true value of the forward-looking parameter. Using Monte-Carlo simulations, we illustrate that the discrepancy of estimators carries on to more complex models, which are non-tractable analytically. Our results shed some light on the long-lasting empirical debate over the importance of the forward-looking component in the hybrid Phillips curve. They suggest that the conflict between estimates reported in the empirical literature may be rationalized by an omitted dynamics.

The issue we are concerned with has been tackled by several recent papers which analyze GMM estimators in the context of the new Phillips curve, e.g. Rudd and Whelan (2001) and Mavroeidis (2001). While our results are consistent with their findings, we extend these studies by considering also the ML estimator. Following Rudd and Whelan (2001), we focus on mis-specification, but the DGP as well as the estimated equation we consider are more general in several respects. Lindé (2001) and Fuhrer and Rudebusch (2002) have compared GMM and ML estimates of hybrid equations using Monte-Carlo simulations. The distinctive feature of the present paper is that we also provide some analytical results.

Section 2 describes the stylized DGP used for our analysis, and describes the GMM and ML estimators. It also investigates the size of the finite-sample bias of those estimators. Section 3 explores the consequences on estimator biases of a measurement error in the forcing variable and of omitted dynamics. Several analytical results concerning the ranking of estimators are proposed. In Section 4, we consider a more general model with some feedback from the dependent variable toward the forcing variable. This case is investigated using Monte-Carlo experiments, since this model cannot be studied analytically. In Section 5, we illustrate that the contrasting results found in the empirical literature on the new Phillips curve can be rationalized using the results obtained in this paper. Section 6 concludes. The Appendix provides further details on the GMM and ML estimators used in the paper and on our analytical results.

2 A stylized DGP with a single lag hybrid equation

2.1 The DGP

In this section, we describe the stylized DGP which includes the hybrid equation and an AR(1) forcing variable and we introduce the estimators that will be used in the remaining of the paper. Let us assume that the DGP is described by a hybrid model in which both a lag and an expected lead of the dependent variable are introduced. The forcing variable is assumed to be an AR(1) process:

$$Y_t = \omega_f E_t Y_{t+1} + (1-\omega_f) Y_{t-1} + \beta Z_t + \varepsilon_t \qquad (2)$$

$$Z_t = \rho Z_{t-1} + u_t \qquad (3)$$

where Y_t is the variable of interest, and Z_t is the true forcing variable. Typically in the hybrid Phillips curve, Y_t represents inflation and Z_t either the output gap or the (log) real unit labor cost. Innovations ε_t and u_t are contemporaneously and serially uncorrelated. We denote $\sigma_\varepsilon^2 = E(\varepsilon_t^2)$ and $\sigma_u^2 = E(u_t^2)$. Structural parameters are therefore $\{\omega_f, \beta, \rho, \sigma_\varepsilon^2, \sigma_u^2\}$.[2]

In order to obtain analytical results, we adopt for the time being a very simple specification for the dynamic of the forcing variable. In addition, we impose that the forward-looking and backward-looking parameters sum to one. This restriction is made for tractability purposes, but also provides an identifying assumption that is necessary in our setup given the process for the forcing variable (see Mavroeidis, 2001, for a discussion of identification). Note that in many instances this restriction is either imposed by the theoretical model or be nearly fulfilled empirically.

As described in Appendix 1, the reduced form of this DGP is obtained by inverting the characteristic polynomial of equation (2), so that

$$(1 - \varphi_1 L)(\varphi_2 - F) Y_t = \frac{\beta}{\omega_f} Z_t + \frac{1}{\omega_f}\varepsilon_t,$$

where L and F denote the lag and forward operators respectively. This leads to the closed-form solution

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

with $\varphi_1 = (1-\omega_f)/\omega_f$, $\varphi_2 = 1$, $\theta = \beta/(\omega_f(1-\rho))$, and $\tilde{\varepsilon}_t = \varepsilon_t/\omega_f$.[3]

Stationarity of the DGP requires that $|\varphi_1| < 1$ (or, equivalently, $0.5 < \omega_f < 1$) and $|\rho| < 1$. These conditions can be weakened by allowing the forcing variable to depend on Y_t. This issue is addressed in details in Section 4.

2.2 Estimators

Since one of the regressors (the expected term) is endogenous, OLS estimation of equation (2) does not yield consistent estimators. Two alternative approaches can then be considered, GMM and ML.

2.2.1 The GMM estimator

The GMM approach reduces to two-stage least-square estimation in this setup. It consists in regressing Y_{t+1} on instruments which are uncorrelated with the error term ε_t but correlated with Y_{t+1}. Since two parameters (φ_1 and θ or, equivalently, ω_f and β) have to be estimated, at least two instruments are needed to achieve identification. Assuming the econometrician is aware of the true specification (2), but does not want to specify the dynamics of the forcing variable, the optimal GMM estimator is obtained using as instrument set $\{Y_{t-1}, Z_t\}$. This estimator relies on the following moment conditions:

$$E\left[Y_{t-1}\left(Y_t - \omega_f Y_{t+1} - (1-\omega_f) Y_{t-1} - \beta Z_t\right)\right] = 0$$
$$E\left[Z_t\left(Y_t - \omega_f Y_{t+1} - (1-\omega_f) Y_{t-1} - \beta Z_t\right)\right] = 0.$$

[2] These parameters are considered as structural here, but they are usually defined as functions of “deep” parameters which reflect constraints and preferences of agents. In the hybrid Phillips curve of Galí and Gertler (1999), the deep parameters include the duration of contracts, the share of rule-of-thumb price setters, and the discount factor.

[3] Note that, without restriction, a more general model could be considered assuming that Z_t is a linear function of q exogenous variables. See Tinsley (2000) for details on this model.

Since the model is just-identified, the probability limits of the estimator are directly obtained by solving these moment conditions.

Notice that, since Z_t is in the instrument set, solving moment conditions is equivalent to the two-step following problem: First, Y_{t+1} is regressed on Y_{t-1} and Z_t giving the fitted value $\hat{Y}_{t+1}$. Second, Y_t is regressed on $\hat{Y}_{t+1}$ and Z_t, yielding consistent estimators.

2.2.2 The ML estimator

The second approach is the ML procedure, which consists in using equation (3) to solve equation (2) iteratively forward. Estimating the reduced form (3)–(4) and imposing cross-equation restrictions allows one to recover the structural parameters. Parameters φ_1 and θ can be estimated directly by OLS. Then, we obtain the estimators of ω_f and β using the relationships $\omega_f = 1/(1+\varphi_1)$ and $\beta = \theta\omega_f(1-\rho)$. Since the estimation is performed after the hybrid model has been solved forward iteratively, ML estimators are obtained under the assumption that forecasts are fully model-consistent. The crucial difference between GMM and ML approaches is that ML imposes some constraints in the way Y_{t+1} is projected onto the state variable, through the dynamics of Z_t used to solve the model. In contrast, GMM does not impose any constraint of this type on the first-stage regression.

Since innovations are assumed to be uncorrelated, this two-step approach is equivalent to the full-information ML. However, when the model is more complex (for instance, when the dynamics of Z_t includes a feedback effect of Y_t), the model generally cannot be solved analytically.

The GMM and ML estimators presented above have the same probability limit when the model estimated is correctly specified. Yet, in many empirical applications of the hybrid model, the gap between GMM and ML estimators has been found to be very large. Two reasons are likely to explain such a discrepancy between GMM and ML estimators: (1) There are differences in the finite-sample properties of the estimators. That the GMM estimator is prone to suffer from finite-sample bias has been underlined by an abundant literature. The size of this bias is related to weak instrument relevance (i.e., weak correlation between instruments and endogenous regressors). This issue has been addressed, among others, by Nelson and Startz (1990), Hall, Rudebusch, and Wilcox (1996), or Staiger and Stock (1997). Other types of finite-sample bias, not related to instruments, may be present in our setup. In particular, in a partial-adjustment context, the autoregressive parameter is biased downward, even when the model is correctly specified (Sawa, 1978). (2) The model is mis-specified, yielding inconsistency of GMM as well as ML estimators. The remaining of the paper investigates these two explanations in turn.

2.3 Finite-sample biases

An abundant literature has studied the finite-sample properties of the GMM estimators, in very different contexts (see Fuhrer, Moore, and Schuh, 1995, or the 1996 special issue of the JBES). These papers provided evidence that GMM estimators may be strongly biased and widely dispersed in finite samples. As far as the ML estimator is concerned, it is in most instances found to be unbiased when the model is correctly specified. Surprisingly, the bias on the autoregressive parameter highlighted by Sawa (1978) is rarely reported in Monte-Carlo experiments.

This section investigates the finite-sample bias of the GMM and ML estimators assuming that the model is correctly specified. As in earlier studies, we rely on Monte-Carlo simulations for this purpose, since the analytical approach in not tractable in this context.

When the GMM approach is implemented, we have to select the instrument set. The optimal instrument set in our setup is $W_t = \{Y_{t-1}, Z_t\}$. This is our baseline case in the following experiments. We also consider the case where lags of W_t are introduced in the instrument set, although they are known to be irrelevant. Thus, we explore instrument sets which include W_t, ..., W_{t-L}, with $L = 0$ and 7.[4] It is worth emphasizing that it is a common practice in the empirical GMM literature to include several lags, without necessarily checking their relevance. A dramatic drawback of this practice is the so-called instrument irrelevance issue, studied in particular by Hall, Rudebusch, and Wilcox (1996), and Staiger and Stock (1997). It can be shown that, when the number of instruments increases in proportion of the number of observations, the probability limit of the GMM estimator tends toward the OLS estimator (Staiger and Stock, 2001, Woglom, 2001). In addition, this bias increases when instrument relevance worsens. We thus have also considered the Plim of the (mis-specified) OLS estimator (see Appendix 1).

2.3.1 Experiment design

To compute the finite-sample distribution of estimators, we rely on Monte-Carlo experiments. Essentially, four structural parameters are likely to affect the finite-sample bias: The forward-looking parameter ω_f, the serial correlation of the forcing variable ρ, the parameter of the forcing variable β, and the signal-to-noise ratio $\sigma_u/\sigma_\varepsilon$. We consider the following parameter sets: $\omega_f = \{0.55; 0.75; 0.95\}$,[5] $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, and $\sigma_u = \sigma_\varepsilon = 1$. Since we found that altering the signal-to-noise ratio within the range $\sigma_u/\sigma_\varepsilon = \{0.5; 2\}$ does not affect finite-sample biases significantly, we maintained in the table the case $\sigma_u = \sigma_\varepsilon = 1$.

The Monte-Carlo experiment is performed as follows. For each parameter set, we simulate $N = 2000$ samples of size $T = 100$. For each simulated sample, a sequence of $T + 100$ random innovations are drawn from the Gaussian distribution $N(0, \Sigma)$ with no serial correlation ($\Sigma = diag(\sigma_u, \sigma_\varepsilon)$), and the first 100 entries are discarded to reduce the effect of initial conditions on the solution path.

2.3.2 Results

Table 1 reports the median and the median of absolute deviations (MAD) of the empirical distribution. It also presents the Plim of the OLS estimator, which is known to be the Plim of GMM estimator in case of instrument irrelevance. The main results are as follows.

First, the ML estimator of ω_f is essentially unbiased, with a very low standard deviation. The estimator of β is unbiased for small values of ρ, while there is a slight positive bias when ρ is large. The median estimator ranges between 0.107 and 0.112, whatever the true parameter ω_f. This bias on β is related to the downward bias of the autoregressive parameter established analytically by Sawa (1978). Since $\beta = \theta\omega_f(1 - \rho)$, the negative bias on ρ translates in a positive bias on β. Note that, since ω_f does not depend on ρ, we do not observe such a bias for ω_f.

[4] In our applications, the long-run covariance matrix is computed using a fixed bandwidth with one lag, as suggested in Newey and West (1987).

[5] We recall that, when the forcing variable is exogenous, stationarity of the model requires $\omega_f > 0.5$.

Second, when the instrument set is correctly chosen ($L = 0$), the GMM(0) estimator of ω_f is slightly biased toward zero, yet not significantly, for large correlation ρ. Moreover, the standard deviation of the estimator is much larger than the standard deviation of the ML estimator. The bias on β is larger and significantly positive for some experiments.

Last, when the instrument set includes irrelevant lags ($L = 7$), the bias on the GMM(7) estimator of ω_f is significantly negative. The higher the true value of ω_f, the larger the bias on the estimator. For $\omega_f = 0.95$, the absolute finite-sample bias is as high as 0.1. In fact, the median estimate lies between the true value of the parameter and the Plim of the OLS, which is found to be close to 0.5 for a wide range of structural parameters. For instance, in the case $\omega_f = 0.75$, $\rho = 0.5$, we obtain $\omega_f = 0.684$ for GMM(7), while Plim $\omega_{OLS} = 0.499$. For large ρ, the bias on β is as high as 30% of the true value.

To sum up, the ML estimators does not display significant finite-sample bias when the model is correctly specified. The GMM estimator of ω_f is not biased significantly, yet the estimator of β displays a positive bias, even when instruments are optimally chosen. When an excessive number of instruments is selected, the estimator of ω_f is more importantly biased toward the OLS estimator. This result confirms previous evidence obtained for instance by Mavroeidis (2001) and Lindé (2001). It is worth noting, however, that the finite-sample bias on parameter ω_f is not sufficient to reconcile the empirical evidence obtained for the hybrid Phillips curve. This is because GMM estimators are systematically biased towards (around) 0.5 in our setup, while estimates reported, for instance, by Galí and Gertler (1999) suggest a bias toward one. In addition, the ML estimator does not display any substantial finite-sample bias. Therefore, if ML estimates were to be close to the true values of the DGP parameters, GMM estimators would display a finite-sample bias toward 0.5, which is not consistent with empirical evidence.

3 Asymptotic biases in mis-specified models

The previous section suggests that finite-sample biases play a limited role in understanding conflicts between estimators. This section now investigates the consequences of two mis-specifications, measurement error and omitted dynamics. Both types of mis-specification appear to be plausible in many applications of the hybrid model. For instance, the new Phillips curve typically includes either the output gap or real marginal cost. The first variable is notoriously difficult to estimate and a wide-range of measures have been proposed in the literature. Real marginal cost can also be measured in a variety of fashion according to the assumption made about the production function. Omitted dynamics is also plausible, since the autoregressive component embodies the behavior of backward-looking price setters. There is no reason for excluding their behavior to depend on several lags of inflation. Alternatively in the consumption function context, general forms of habit formation (such as in Otrok, Ravikumar, and Whiteman, 2002) can lead to several lags of consumption in the reduced form.

3.1 Mis-specification of type I: Measurement error

3.1.1 Analytical results

In order to illustrate the case of measurement error, we now adopt the following DGP:

$$Y_t = \omega_f E_t Y_{t+1} + (1 - \omega_f) Y_{t-1} + \beta Z_t + \varepsilon_t$$

$$
\begin{aligned}
Z_t &= \rho Z_{t-1} + u_t \\
X_t &= a Z_t + e_t
\end{aligned} \tag{5}
$$

where X_t is the proxy of the forcing variable used by the econometrician (e.g., the output gap or the real ULC in the context of the hybrid Phillips curve). Innovation e_t is assumed to be contemporaneously and serially uncorrelated with ε_t and u_t. We denote $\sigma_e^2 = E\left(e_t^2\right)$. We also define $\tau = a^2\sigma_Z^2 / \left(a^2\sigma_Z^2 + \sigma_e^2\right)$, the fraction of the variance of X_t explained by Z_t. Parameter τ can be viewed as a measure of the quality of the proxy X_t. When τ gets closer to one, the quality of the proxy improves. Although we expect parameter a to be positive, it may be negative as well.

Since the econometrician is assumed to erroneously select X_t as the forcing variable, the following mis-specified model is estimated:

$$
Y_t = \alpha_f E_t Y_{t+1} + (1-\alpha_f) Y_{t-1} + b X_t + \upsilon_t \tag{6}
$$

$$
X_t = \psi X_{t-1} + w_t. \tag{7}
$$

The estimated degree of forward-lookingness α_f is a presumably biased estimator of the parameter ω_f. Note that there is no mis-specification in the limiting case where $\sigma_e^2 = 0$, i.e. $\tau = 1$, so that X_t is actually the true forcing variable.

The reduced form of the postulated system (6)–(7) is given by

$$
Y_t = \varphi Y_{t-1} + \mu X_t + \tilde{\varepsilon}_t. \tag{8}
$$

Stationarity conditions are, as previously, $\varphi < 1$ (or $\omega_f > 0.5$) and $|\psi| < 1$.[6]

Since two parameters (α_f and b) have to be estimated in the mis-specified model, at least two instruments are needed to achieve identification. Two cases appeared to us of particular interest. The first estimator (GMM1) is based on the instrument set $\{Y_{t-1}, X_t\}$, while the second estimator (GMM2) resort to the instrument set $\{Y_{t-1}, X_t, Z_t\}$. In the latter case, the instruments include the actual forcing variable. Both cases are likely to occur in empirical applications. For instance, the first estimator may correspond to the case of the hybrid I-S curve, as studied recently by Fuhrer and Rudebusch (2002). Instruments are lagged output gap, inflation rate, and interest rate. Mis-specification is likely to occur because the definition of the real interest rate may not be relevant. In contrast, in the case of the hybrid Phillips curve, the theoretically relevant forcing variable is the real marginal cost, while estimations are performed using, alternatively, the output gap and the real ULC. Most studies (in the following of Galí and Gertler, 1999) therefore include both the output gap and real ULC in the instrument set. We may expect the real marginal cost to be well proxied by a linear combination of the two variables.

Estimator GMM1 relies on the following moment conditions:

$$
\begin{aligned}
E\left[Y_{t-1}\left(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f) Y_{t-1} - b X_t\right)\right] = 0 \\
E\left[X_t\left(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f) Y_{t-1} - b X_t\right)\right] = 0.
\end{aligned}
$$

[6] Combining equations (3) and (5) should yield an ARMA process for X_t. This is because the measurement error introduces a serial correlation of order 1 in the error term of X_t. The process for X_t can be written as $X_t = \rho X_{t-1} + w_t - c w_{t-1}$, with $c = \left(\sigma_w^2 \pm \sqrt{\sigma_w^4 - 4\rho^2\sigma_e^4}\right) / \left(2\rho\sigma_e^2\right)$ and $\sigma_w^2 = \rho\sigma_e^2 / c$. Such an approach is not followed, however, because it would not allow to obtain analytical results. In addition, it would not be consistent with the usual approach of the econometrician.

Since the model is just-identified, the probability limits of the estimators are directly obtained by solving the two moment conditions.

In contrast, since estimator GMM2 corresponds to an over-identified case, this estimator cannot be obtained in a trivial way. It proves convenient to view this estimator as a two-step estimator: First, we regress Y_{t+1} on the instrument set to build the fitted value $\hat{Y}_{t+1}$. Then, we estimate, by OLS:

$$Y_t - Y_{t-1} = \alpha_f \left(\hat{Y}_{t+1} - Y_{t-1} \right) + bX_t + \xi_t.$$

The ML estimator is obtained by estimating the reduced form of the model, given by equation (8). Parameters φ and μ can be estimated by OLS. Then, we obtain the estimators of α_f and b with the following relationships: $\alpha_f = 1/\left(1+\varphi\right)$, $b = \mu\alpha_f(1-\psi)$.

Since the hybrid model is estimated using X_t as forcing variable in place of Z_t, there exists an asymptotic bias, the extent of which is directly related to the correlation between the two variables X_t and Z_t. The following proposition gives the asymptotic biases of GMM as well as ML estimators. Details on the calculation of these biases are relegated in the Appendix 1.

Proposition 1: Assume that the DGP is given by equations (2)–(3). Assume that the econometrician estimates the model with X_t in place of Z_t as the forcing variable, corresponding to equations (6)–(7). Then, the three estimators defined above have the following probability limits:

- GMM estimator with instrument set $\{Y_{t-1}, X_t\}$ (GMM1):

$$\begin{aligned}
\alpha_{GMM1} &= \omega_f \left(\frac{\sigma_\varepsilon^2 + \Lambda^2 \left[1 - (1-\tau)\frac{\rho(1+\varphi_1)}{1+\rho} \right]}{\sigma_\varepsilon^2 + \Lambda^2 \left[1 - (1-\tau)\rho\varphi_1 \right]} \right) \\
b_{GMM1} &= \left(\frac{\beta}{a}\tau \right) \left(\frac{\sigma_\varepsilon^2 + \Lambda^2}{\sigma_\varepsilon^2 + \Lambda^2 \left[1 - (1-\tau)\rho\varphi_1 \right]} \right)
\end{aligned}$$

- GMM estimator with instrument set $\{Y_{t-1}, X_t, Z_t\}$ (GMM2):

$$\begin{aligned}
\alpha_{GMM2} &= \omega_f \left(\frac{\sigma_\varepsilon^2 + \Lambda^2 \left[1 + (1-\tau)\frac{(1-\rho)\varphi_1}{1-\varphi_1} \right]}{\sigma_\varepsilon^2 + \Lambda^2 \left[1 + (1-\tau)\frac{(1-\rho^2)\varphi_1^2}{(1-\varphi_1^2)} \right]} \right) \\
b_{GMM2} &= \left(\frac{\beta}{a}\tau \right) \left(\frac{\sigma_\varepsilon^2 + \Lambda^2}{\sigma_\varepsilon^2 + \Lambda^2 \left[1 + (1-\tau)\frac{(1-\rho^2)\varphi_1^2}{(1-\varphi_1^2)} \right]} \right)
\end{aligned}$$

- ML estimator:

$$\begin{aligned}
\alpha_{ML} &= \omega_f \left(\frac{\sigma_\varepsilon^2 + \Lambda^2 \left[1 + (1-\tau)\frac{(1-\varphi_1^2)\rho^2}{1-\rho^2} \right]}{\sigma_\varepsilon^2 + \Lambda^2 \left[1 + (1-\tau)\frac{(1-\varphi_1)\rho}{1-\rho} \right]} \right) \\
b_{ML} &= \left(\frac{\beta}{a}\tau \right) \frac{1-\psi}{1-\rho} \left(\frac{\sigma_\varepsilon^2 + \Lambda^2}{\sigma_\varepsilon^2 + \Lambda^2 \left[1 + (1-\tau)\frac{(1-\varphi_1)\rho}{1-\rho} \right]} \right)
\end{aligned}$$

where $\Lambda = \beta\sigma_u \left(1+\varphi_1\right) / \left[\left(1-\varphi_1\rho\right)\left(1-\rho\right) \right]$ and $\psi = \rho\tau$ is the Plim of the autoregressive parameter of X_t. Parameter Λ^2 can be interpreted as the fraction of the variance of Y_t explained by Z_t.

We verify that, in the case $\tau = 1$, all estimators α_f are asymptotically unbiased, so that $\alpha_{GMM1} = \alpha_{GMM2} = \alpha_{ML} = \omega_f$. Similarly, all estimators b are asymptotically unbiased, since $b_{GMM1} = b_{GMM2} = b_{ML} = \beta/a$. In the case $\rho = 0$, $\alpha_{GMM1} = \alpha_{ML} = \omega_f$, while $\alpha_{GMM2} > \omega_f$.

Corollary 1: Under the condition $\varphi_1 \rho < 1$, the following inequalities hold:

$$\begin{aligned} 1/4 &< \alpha_{GMM1} < \omega_f \\ 1/2 &< \alpha_{ML} < \omega_f < \alpha_{GMM2} < 1 \end{aligned}$$

and when $\rho - \varphi_1 < (1 - \varphi_1^2)(1 - \tau)\,\rho\Lambda/\sigma^2$:

$$\alpha_{GMM1} < \alpha_{ML}.$$

In addition, the following inequalities hold:

$$b_{GMM2} < (\tau\beta/a) < b_{GMM1} < b_{ML}.$$

These inequalities are readily obtained by comparing terms between brackets in the numerator and denominator of each expression. Bounds for estimators of α_f are obtained, under the stationarity assumption, by choosing bounds for structural parameters ω_f (1/3 and 1), ρ (0 and 1), and $\tau = \sigma_\varepsilon = 0$. We thus obtain that estimators ML and GMM2 of α_f are biased *in opposite directions* from the true value of the parameter ω_f.

3.1.2 Evidence and discussion

Table 2 presents probability limits of GMM and ML estimators in the measurement-error case, using formulae reported in proposition 1. We select as structural parameters $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, and $\sigma_u/\sigma_\varepsilon = 1$. Last, we choose $\tau = \{0.1; 0.5; 0.9\}$, which represents to which extent the variable X_t is a good proxy for Z_t. Parameter τ plays presumably a crucial role regarding asymptotic biases.

As regards parameter α_f, illustrating proposition 1, the estimators GMM1 and ML are systematically lower than the true value ω_f, while the estimator GMM2 is systematically larger than ω_f. Since the estimators are bounded by 0.5, biases on GMM1 and ML are, as expected, very small when ω_f is equal to 0.55 and, conversely, the bias on GMM2 is very small when ω_f is equal to 0.95.

Unsurprisingly, the lower the quality of the proxy (τ), the larger the bias on α_f. An interesting result is that, for low values of τ and ρ, the GMM2 bias on α_f can be extremely large. For instance, for $\tau = \rho = 0.1$ and $\beta = 1$, the probability limit is as high as 0.89, so that the dependent variable Y_t would be claimed to be an essentially forward-looking process, while the true weight on the forward-looking component is in fact 0.55. Conversely, the ML estimator can be severely biased for low value of τ and large value of ρ. For $\tau = 0.1$, $\rho = 0.9$, and $\beta = 1$, the Plim is 0.53, so that the forward-looking weight would be claimed to be at its lower bound, while Y_t is in fact essentially forward-looking.

In addition, the ranking of GMM1 and ML depends on values of τ and ρ. When ω_f is small, only very large values of τ and ρ yield a smaller bias on GMM1 than on ML. In contrast, when ω_f is large, the bias on GMM1 is smaller as soon as $\rho \geq 0.5$.

Last, the probability limits reported for the estimator of b are very close to $\beta\tau$ when ω_f is large (0.75 or 0.95). We retrieve the ranking highlighted in corollary 1. Note that the

ML estimator is systematically larger than GMM estimators and for small ω_f it can also be much larger than the true parameter $\beta = 1$. It is worth emphasizing that the estimator of b is not likely to be negative in the general case. Smallest values of b are obtained for the GMM2 estimator when ω_f, ρ and τ are low.

To sum up, GMM estimation can lead to overestimation of the degree of forward-lookingness. The bias is potentially large. This finding echoes Rudd and Whelan (2001), although these authors considered omitted variables rather than measurement errors. Note that our setup is in some respect more general since both the estimated equation and the DGP are hybrid models, while Rudd and Whelan considered the case of a purely backward-looking DGP and a purely forward-looking estimated equation. The crucial feature in the two cases is that a positive GMM bias occurs when the relevant forcing variable is not introduced as regressor in the model, but *is included in the instrument set*. The mechanics is that actual future inflation used in the first-stage regression captures the effect of the (omitted) relevant variable, and hence the second-stage regression will tend to put an excessive weight on the fitted value. Since we find that the ML bias has a negative sign, if any, this may be a step towards rationalizing empirical conflicts.

3.2 Mis-specification of type II: Omitted dynamics

3.2.1 Analytical results

We now consider another plausible mis-specification, the case of omitted dynamics. We assume that the true DGP includes two lags of the dependent variable:

$$\begin{aligned} Y_t &= \omega_f E_t Y_{t+1} + \omega_b^1 Y_{t-1} + \left(1 - \omega_f - \omega_b^1\right) Y_{t-2} + \beta Z_t + \varepsilon_t \\ Z_t &= \rho Z_{t-1} + u_t. \end{aligned} \tag{9}$$

The reduced form of this model is thus

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

where parameters φ_1 and φ_2 are related to the roots of the characteristic polynomial:

$$\left(-\omega_f + L - \omega_b^1 L^2 - \left(1 - \omega_f - \omega_b^1\right) L^3\right) F Y_t = \beta Z_t + \varepsilon_t.$$

They are given by $\varphi_1 = (1 - \omega_f)/\omega_f$, $\varphi_2 = (1 - \omega_f - \omega_b^1)/\omega_f$, and $\theta = \beta/\left(\omega_f(1-\rho)\right)$. Further details on the model with two lags are provided in Appendix 2. In particular, stationarity conditions of equations (3) and (9) are $\left(2 - 3\omega_f - \omega_b^1\right) < 0$, $-\omega_f < \omega_b^1 < 1$, and $|\rho| < 1$.

Now, the estimated mis-specified model is a one-lag hybrid model:

$$Y_t = \alpha_f E_t Y_{t+1} + (1 - \alpha_f) Y_{t-1} + b Z_t + v_t. \tag{10}$$

The difference with the previous mis-specification is that the econometricain uses the true forcing variable, but omits the presence of the second lag. As in the case of measurement error, we consider two instrument sets for GMM estimators. GMM as well as ML estimators can be built quite similarly to the case with measurement errors.

Estimator GMM1 includes $\{Y_{t-1}, Z_t\}$ as instruments, leading to a just-identified parameter set. It relies on the following moment conditions:

$$\begin{aligned} E\left[Y_{t-1}\left(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f) Y_{t-1} - b Z_t\right)\right] &= 0 \\ E\left[Z_t\left(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f) Y_{t-1} - b Z_t\right)\right] &= 0. \end{aligned}$$

Since the model is just-identified, the probability limits of the estimators are directly obtained by solving the two moment conditions.

GMM2 includes $\{Y_{t-1}, Y_{t-2}, Z_t\}$ leading to an over-identified parameter set, with the property that the omitted variable appears in the instrument set. Note that GMM2 mimics actual practice of GMM of introducing several lags of the relevant variables in the instrument set. Therefore, GMM2 involves a third moment conditions, and thus corresponds to an over-identified case. As in the case of measurement error, this estimator is built as a two-step estimator.

The ML estimator is obtained by estimating the reduced form of the postulated system (3)–(10), that is:

$$Y_t = \varphi Y_{t-1} + \mu Z_t + \tilde{\varepsilon}_t.$$

Parameters φ and μ are estimated by OLS. Then, estimators of α_f and b are defined by the following relationships $\alpha_f = 1/(1+\varphi)$ and $b = \mu\alpha_f(1-\rho)$.

The probability limit of each estimator is given in Proposition 2. Details on the computation of the probability limits are provided in Appendix 2.

Proposition 2: Assume that the DGP is given by equations (3) and (9). Assume that the econometrician estimates the model omitting the second lag of Y_t, corresponding to the model given by equations (3) and (10). Then, the three estimators have the following probability limits:

- GMM estimator with instrument set $\{Y_{t-1}, Z_t\}$ (GMM1):

$$\begin{aligned}
\alpha_{GMM1} &= \left(\frac{1}{1+\varphi_1-\varphi_2}\right)\left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\left(-1+\varphi_1-\varphi_2+\varphi_2\rho\right)\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\left(\varphi_1+\varphi_2\rho\right)\right]}\right) \\
b_{GMM1} &= \beta\left(\frac{(1+\varphi_1)\left(1-\varphi_1\rho-\varphi_2\rho^2-\varphi_2(1+\rho)\right)}{(1+\varphi_1-\varphi_2)\left(1-\varphi_1\rho-\varphi_2\rho^2\right)}\right) \\
&\times\left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\left(\varphi_1+\varphi_2\rho\right)\frac{(1-\varphi_1\rho-\varphi_2\rho^2)+(1+\rho)}{(1-\varphi_1\rho-\varphi_2\rho^2)-\varphi_2(1+\rho)}\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\left(\varphi_1+\varphi_2\rho\right)\right]}\right)
\end{aligned}$$

- GMM estimator with the instrument set $\{Y_{t-1}, Y_{t-2}, Z_t\}$ (GMM2):

$$\begin{aligned}
\alpha_{GMM2} &= \left(\frac{1-\varphi_1-\varphi_1\varphi_2}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right)\left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{-1+\varphi_1-\varphi_2+\varphi_2\rho}{1-\varphi_1-\varphi_1\varphi_2}\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{-2\varphi_1\varphi_2+\varphi_2\rho-\varphi_2^2\rho}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right]}\right) \\
b_{GMM2} &= \beta\left(\frac{(1+\varphi_1)\left(1-\varphi_1-\varphi_2-\rho\left(1-\varphi_1-\varphi_1\varphi_2\right)\left(\varphi_1+\varphi_2+\varphi_2\rho\right)\right)}{\left(1-\varphi_1\rho-\varphi_2\rho^2\right)\left(1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2\right)}\right) \\
&\times\left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{2\varphi_1+\varphi_2\rho-\varphi_2^2\rho-(\varphi_1+\varphi_2\rho)(-1+\varphi_1+\varphi_2-\rho(1-\varphi_1+\varphi_2\rho))}{1-\varphi_1-\varphi_2-\rho(1-\varphi_1-\varphi_1\varphi_2)(\varphi_1+\varphi_2+\varphi_2\rho)}\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{2\varphi_1+\varphi_2\rho-\varphi_2^2\rho}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right]}\right)
\end{aligned}$$

- ML estimator:

$$\begin{aligned}
\alpha_{ML} &= \left(\frac{1-\varphi_2}{1+\varphi_1-\varphi_2}\right)\left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{2\varphi_1+\varphi_2\rho-\varphi_2^2\rho}{1-\varphi_2}\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{2\varphi_1+\varphi_2\rho-\varphi_2^2\rho+1+\varphi_1^2-\varphi_2^2+\varphi_1\varphi_2\rho}{1+\varphi_1-\varphi_2}\right]}\right) \\
b_{ML} &= \beta\left(\frac{(1+\varphi_1)\left(1+\varphi_1-\varphi_2-\rho\left(1-\varphi_2\right)\right)}{\left(1-\varphi_1-\varphi_2\right)\left(1-\varphi_1\rho-\varphi_2\rho^2\right)}\right)
\end{aligned}$$

$$\times\left(\frac{\sigma_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{(1-\rho)\left(2\varphi_1+\varphi_2\rho-\varphi_2^2\rho\right)+1+\varphi_1^2-\varphi_2^2+\varphi_1\varphi_2\rho}{1+\varphi_1-\varphi_2-\rho(1-\varphi_2)}\right]}{\sigma_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{\left(2\varphi_1+\varphi_2\rho-\varphi_2^2\rho\right)+1+\varphi_1^2-\varphi_2^2+\varphi_1\varphi_2\rho}{1-\varphi_1-\varphi_2}\right]}\right)$$

where $\tilde{\Lambda}=\beta\sigma_u\left(1+\varphi_1\right)/\left[\left(1-\varphi_1\rho-\varphi_2\rho^2\right)(1-\rho)\right].$

We verify that, in the case $\varphi_2=0$, all estimators are unbiased, since we have $\alpha_{GMM1}=\alpha_{GMM2}=\alpha_{ML}=\omega_f=1/\left(1+\varphi_1\right)$. Since we are primarily interested in the effect of omitting a lag, we consider more specifically the case $\rho=0$, which corresponds to the forcing variable being a pure white noise. We then obtain the following corollary.

Corollary 2: In the case where $\rho=0$, estimators of α_f are

$$\begin{aligned}
\alpha_{GMM1} &= \left(\frac{1}{1+\varphi_1-\varphi_2}\right)\\
\alpha_{GMM2} &= \left(\frac{1-\varphi_1-\varphi_1\varphi_2}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right)\\
\alpha_{ML} &= \left(\frac{1-\varphi_2}{1+\varphi_1-\varphi_2}\right).
\end{aligned}$$

Under the additional condition $\varphi_2>0$ (or, equivalently, $\omega_b^2>0$), the following inequalities hold:

$$\alpha_{ML}<\omega_f<\alpha_{GMM1}<\alpha_{GMM2},$$

while under the additional condition $\varphi_2<0$ (or, equivalently, $\omega_b^2<0$), we have:

$$\begin{aligned}
&\alpha_{ML}<\omega_f<\alpha_{GMM2}<\alpha_{GMM1} && \text{when } \varphi_1>1\\
&\alpha_{GMM1}<\omega_f<\alpha_{GMM2}<\alpha_{ML} && \text{when } \varphi_1<1.
\end{aligned}$$

In the case when $\rho=0$, estimators of b are

$$\begin{aligned}
b_{GMM1} &= \beta\left(\frac{\left(1+\varphi_1\right)\left(1-\varphi_2\right)}{1+\varphi_1-\varphi_2}\right)\\
b_{GMM2} &= \beta\left(\frac{\left(1+\varphi_1\right)\left(1-\varphi_1-\varphi_2\right)}{1-\varphi_1^2-\varphi_1^2\varphi_2-\varphi_2}\right)\\
b_{ML} &= \beta\left(\frac{\left(1+\varphi_1\right)\left(1+\varphi_1-\varphi_2\right)}{1-\varphi_1-\varphi_2}\right)
\end{aligned}$$

so that the following inequalities hold:

$$\begin{aligned}
b_{GMM2} &< b_{GMM1}<\beta<b_{ML} && \text{when } \varphi_2>0\\
b_{GMM2} &< b_{GMM1}<b_{ML}<\beta && \text{when } \varphi_2<0.
\end{aligned}$$

Corollary 3: In the general case, the following inequalities hold between estimators of α_f, when $\varphi_2>0$:

$$\begin{aligned}
\alpha_{ML} &< \left(\frac{1-\varphi_2}{1+\varphi_1-\varphi_2}\right)<\omega_f\\
\alpha_{ML} &< \left(\frac{1-\varphi_2}{1+\varphi_1-\varphi_2}\right)<\alpha_{GMM1},
\end{aligned}$$

while, when $\varphi_2<0$:

$$\begin{aligned}
\omega_f &< \alpha_{ML}\\
\alpha_{GMM1} &< \alpha_{ML}.
\end{aligned}$$

3.2.2 Evidence and discussion

Table 3 presents probability limits of GMM and ML estimators in the omitted dynamics case, using formulae reported in Proposition 2. We select as structural parameters $\omega_f = \{0.4; 0.65; 0.9\}$, $\rho = \{0; 0.5; 0.9\}$, and $\beta = \{0.1; 1\}$. We only report results when $\sigma_u = \sigma_\varepsilon = 1$, because it has been found to have only marginal effects on estimators. Last, we choose (ω_b^1, ω_b^2) such that the persistence of the dependence variable is $\varphi_1 + \varphi_2 = \{0; 0.5; 0.9\}$, corresponding to a low, medium, and strong persistence in the dependent variable dynamics.

The results reported in Table 3 contrast rather sharply with those reported in the case of measurement error. An important difference is that very large biases are likely to occur whatever the true value of ω_f. For instance, when $\omega_f = 0.4$, with a low persistence which corresponds to a large negative ω_b^2, the estimator of α_f is strongly biased toward 0 for GMM1 (0.25 and 0.07 with $\rho = 0$ and 0.9 respectively), but toward 1 for ML (0.63 and 0.52 with $\rho = 0$ and 0.9). In case of a large persistence (while maintaining $\omega_f = 0.4$), which mimics the results of Fuhrer for the new Phillips curve, we obtain limited biases. The Plim of α_f ranges between 0.3 and 0.6, whatever the estimation procedure.

When the forward-looking component is very large ($\omega_f = 0.9$ close to the results of Galí and Gertler), we also obtain that GMM estimators are likely to be severely biased when the missing lag has a large positive parameter. In those cases, α_f has a Plim equal to 1.4 for both GMM procedures (with $\omega_b^2 = 0.35$) and as high as 3 for GMM1 and 4 for GMM2 when $\omega_b^2 = 0.71$. Contrasting with these extreme outcomes, the ML estimator displays only moderate biases.

The effect of the serial correlation of the forcing variable is ambiguous. It is likely to affect the parameter estimate rather strongly but in a way which depends both on the weight of the forward-looking component and on the persistence of the endogenous variable.

In sum, the ML estimators of α_f and b are rather moderately biased even for extreme cases. This result contrasts sharply with our findings for the GMM estimators which are likely to display severe biases in cases of a large forward-looking component.

4 A model with endogenous forcing variable

This section addresses the issue of an endogenous forcing variable. This issue is particularly interesting for several reasons. First, it is likely to be the case that the dynamics of the forcing variable is related to the dependent variable. In the case of the hybrid Phillips curve, the real marginal cost depends on inflation quite naturally (see Sbordone, 2002, or Kurmann, 2002). Also in an inflation/output gap model, we expect some feedback from inflation to the output gap through the reaction function of the central bank. Second, rendering the forcing variable endogenous allows to consider a wider interval for the weight of the forward-looking component. Indeed, when Z_t is exogenous, only values of ω_f larger than 0.5 are consistent with stationarity of the DGP with a single lag. When Z_t is endogenous, virtually all values of ω_f are consistent with stationarity, provided the extent of endogeneity is sufficiently large. The only constraint is that the feedback effect from Y_t to Z_t has a sign opposite to that of β.

4.1 The DGP

The DGP we now consider is

$$
\begin{aligned}
Y_t &= \omega_f E_t Y_{t+1} + (1-\omega_f) Y_{t-1} + \beta Z_t + \varepsilon_t \\
Z_t &= \rho Z_{t-1} + \theta Y_{t-1} + u_t.
\end{aligned}
$$

The case of endogenous forcing variable cannot solved analytically, because it involves a joint estimation of the two equations. Yet, numerical solutions can be quite easily obtained. For GMM, the standard approach applies. For ML, we resort to the procedures proposed, for instance, by Anderson and Moore (1995), Söderlind (1999), or Klein (2000). These procedures are able to solve, numerically, the forward-looking component iteratively. Then, standard ML estimation applies to the purely backward-looking reduced form of the DGP.

The aim of this section is twofold. First, we wish to study the finite-sample biases in a correctly-specified model with endogenous forcing variable. As claimed before, when the forcing variable is exogenous, stationarity of the DGP requires $\omega_f > 0.5$. In this setup, we obtained in Section 2.2 that the estimator of ω_f has a negative bias. Allowing the true parameter ω_f to be lower than 0.5 gives the opportunity to investigate this issue more deeply. As we will show, the finite-sample bias is indeed toward the Plim of the OLS estimator, and turns out to be positive when ω_f is small.

The second issue we wish to address is the asymptotic bias in a mis-specified model. Once again, since the weight of the forward-looking component is constrained to be larger than a given bound (0.5 with a single lag and 0.33 with two lags), we cannot study where estimators α_f are located when ω_f is small. Yet, this is an important issue since several authors have advocated that the forward-looking component is in fact very small (see Fuhrer and Moore, 1995b, Fuhrer, 1997, Fuhrer and Rudebusch, 2002). Again the asymptotic bias is obtained through Monte-Carlo simulations.

4.2 The finite-sample bias in the correctly specified model

4.2.1 Experiment design

As baseline case, we select the following parameter values: $\omega_f = 0.25$, $\rho = 0.75$, and $\sigma_u^2 = \sigma_\varepsilon^2 = 1$. We choose $\omega_f = 0.25$ because the main interest of endogeneizing the forcing variable is to allow for a low value of ω_f. Such a value has been found to be plausible in Fuhrer (1997) in the case of the Phillips curve. We select values of β in $\{0.1; 1\}$ andθ in $\{-0.5; -0.1\}$. The value of $\theta = -0.5$ is taken from Fuhrer and Rudebusch (2002). It is arguably large, but our purpose is to illustrate the effect of endogeneity of the forcing variable. Other values considered are $\omega_f = \{0.5; 0.75\}$ and $\rho = \{0.5; 0.9\}$.[7]

4.2.2 Results

Table 4 reports results for GMM with 0 and 7 lags and for ML. We also report the Plim of OLS (computed from a simulation with $T = 100,000$). For $\omega_f = 0.5$ and 0.75, we find

[7] To insure comparability of estimators, we constraint ω_f to be located between 0 and 1. To implement this constraint with GMM, we estimate unrestrictedly in the first stage. Then, we impose the value of ω_f to be equal to 0 (respectively 1) when the GMM estimates is below 0 (above 1). Then, we re-estimate the parameter of the forcing variable under this constraint. For the ML, we use the Constrained Maximum Likelihood package of GAUSS.

similar results to the case with no feedback, so we focus on the case $\omega_f = 0.25$ only. Our main findings are as follows:

First, the GMM bias negligible for $\beta = 1$ and moderate for $\beta = 0.1$. This is because the bias of GMM is in direction of the OLS bias. When $\beta = 0.1$, the Plim of the OLS estimator is very close to 0.5, so that the median of the GMM estimator ranges between 0.3 and 0.5. Yet, when $\beta = 1$, the Plim of OLS is much closer to 0.25, so that even OLS is nearly unbiased.

Second, in some instances, ML appears to suffer from a noticeable finite-sample bias. This bias is in fact related to the finite-sample bias of the OLS estimate of the autoregressive parameter in an AR(1) model (see, e.g., Sawa, 1978). The bias is substantial only when the autoregressive root of the univariate process for Y_t is large, i.e. when ω_f is small. However, given that the estimated structural parameter is a nonlinear function of the AR parameter, a small bias on the AR root translates into a noticeable bias for ω_f. ML also experiences low performance, when ρ is low, because there is weak empirical identification. Indeed in this case, Z_t can hardly be distinguished from past value of Y_t.

A last point is that the estimator of b is generally biased toward 0 when $\beta = 0.1$, while it is biased upwards when $\beta = 1$. These biases are moderate, however.

To sum up, the main conclusions found previously extend to this more plausible case. With GMM, the finite-sample bias of ω_f is generally positive. It is worth emphasizing, however, that the parameter is biased toward 0.5 and not toward 1. Therefore, it cannot account for GMM estimates reported in the literature.

4.3 The asymptotic bias in the case of mis-specification

Since asymptotic biases cannot be computed analytically in this setup, they are computed using a large Monte-Carlo simulation ($T = 25,000$) for each experiment.

4.3.1 Measurement error

The baseline parameters are, as previously, $\omega_f = 0.25$, $\rho = 0.75$, and $\sigma_u^2 = \sigma_\varepsilon^2 = 1$. We also select values of θ in $\{-0.5; -0.1\}$ and β in $\{0.1; 1\}$. We set $\tau = 0.5$ and $\tau = 0.75$ as alternative values for the quality of the proxy. **Table 5** reports the results of these simulations.

As a preliminary comment, it is worth emphasizing that, unlike the simple model of Section 2, both GMM and ML estimators are now biased *upwards*. This result is expected, since that forward-looking component is chosen to be very low ($\omega_f = 0.25$). The bias on ω_f is moderate when the quality of the proxy, τ, is large, with a bias smaller than 0.1. In contrast, when τ is chosen to be small ($\tau = 0.1$), we obtain severe biases, as high as 0.3 for the GMM as well as the ML estimators. By and large, biases are found to be located in a similar range for the three estimators.

As expected, biases are smaller when the forward-looking parameter ω_f is equal to 0.5 or 0.75. Yet, for a small feedback parameter θ, we obtain large negative biases when $\omega_f = 0.75$.

In many instances, the parameter of the forcing variable (b) is found to be severely biased toward 0. This is likely to rationalize the very small parameters obtained in previous empirical estimates of the new Phillips curve.

Our evidence suggests that such a measurement error, as described in this paper, is very unlikely to fill the gap between parameter estimates obtained by GMM and ML, even in the case of an endogenous forcing variable.

4.3.2 Omitted dynamics

The baseline parameters are unchanged, except for the weight of the components. We consider the following sets for $(\omega_f, \omega_b^1, \omega_b^2)$: $(1/3, 1/3, 1/3)$, $(1/2, 1/4, 1/4)$, $(1/2, 0, 1/2)$, $(0.65, 0.35, 0)$ [no mis-specification in this case] and $(3/4, 3/4, -1/2)$, $(3/4, -1/2, 3/4)$.

Table 6 reports the results. Interestingly, we notice that biases are very substantial for GMM estimators, even when the weight of the omitted lag is low. An exception is of course the case with no mis-specification. For instance, when $\omega_f = 0.5$, $\omega_b^1 = 0.25$, and $\omega_b^2 = 0.25$, we obtain a Plim of estimator equal to 0.95 for GMM2, and 0.58 for ML. The Plim of GMM estimator is often higher than 1, even when $\omega_f \leq 0.5$, so that the econometrician would reject the hybrid model in favor of a purely forward-looking model. In addition, biases are particularly large for the GMM2 estimator. The Plim of the estimator α_f is found to be larger than 1 in all instances but two cases, corresponding to the lowest weights on the second lag. The result that the bias for both GMM and ML estimators is positive when ω_b^2 is positive provides to some extent some generality to Corollary 3. However, although the ML bias is generally positive, it is much smaller than the GMM bias.

When the feedback parameter θ decreases in absolute value, the bias on GMM estimators increases systematically. The case $\beta = 1$ helps to reduce persistence in the system, so that we obtain much lower biases than for $\beta = 0.1$.

A further interesting result is that the Plim of the GMM estimator of b is found in many instances to be negative. This is the case in particular when the weight ω_b^2 is positive and large and when the parameter of the forcing variable is small ($\beta = 0.1$). Also the bias on b is much larger for GMM2 than for GMM1.

On the whole, these experiments show that a limited amount of mis-specification in the form of an omitted dynamics may produce discrepancies between GMM and ML estimators of the degree of forward-lookingness that are in excess of 0.5.

5 Rationalizing evidence on the hybrid Phillips curve

5.1 The hybrid Phillips curve

The basic hybrid inflation model is the following:

$$\pi_t = \omega_f E_t \pi_{t+1} + (1 - \omega_f)\, \pi_{t-1} + \beta Z_t + \varepsilon_t \tag{11}$$

where π_t is the inflation rate and Z_t is the forcing variable. The parameter ω_f is the weight on the forward-looking component, with $0 \leq \omega_f \leq 1$. Such a model has been proposed originally by Chadha, Masson, and Meredith (1992). We consider two cases for Z_t: the output gap $(\widehat{y}_t)$ and the marginal cost $(\widehat{mc}_t)$, usually proxied by the deviation of real ULC from steady state.[8] With the output gap $\widehat{y}_t$ as a forcing variable, this model nests as special cases the traditional Phillips curve $(\omega_f = 0)$ as well as the Taylor (1980) forward-looking

[8] The output gap itself may be viewed as a proxy of real marginal cost, yet under more restrictive assumptions.

Phillips curve ($\omega_f = 1$). It also nests the Fuhrer and Moore (1995b) model with two-period contracts ($\omega_f = 1/2$). With marginal cost $\widehat{mc}_t$ as a forcing variable, this model is very close to the hybrid Phillips curve put forward by Galí and Gertler (1999). One important feature of this latter model is that it has some micro-foundations. Galí and Gertler have derived the hybrid Phillips curve assuming that some of the firms set their price optimally in sticky-price framework. As in the baseline Calvo (1983) model, only a fraction of firms is allowed to reset their price, at each date. Among the firms which are able to change their price, some are forward-looking and some are backward-looking. Galí and Gertler then show that the equation for aggregate inflation has the following form:

$$\pi_t = (\phi^{-1}\omega)\pi_{t-1} + (\phi^{-1}\delta\alpha)E_t\pi_{t+1} + (1-\alpha)(1-\omega)(1-\alpha\delta)\phi^{-1}Z_t$$

where δ is the given discount rate, $(1-\alpha)$ is the probability of each firm to be allowed to reset its price, ω is the share of rule-of-thumb price setters, and $\phi = [\alpha + (1-\alpha)\omega + \omega\alpha\delta]$. Given that all parameters are not identifiable using a single-equation approach, an empirically equivalent specification is given by:

$$\pi_t = \omega_f E_t\pi_{t+1} + \omega_b\pi_{t-1} + \beta Z_t + \varepsilon_t. \tag{12}$$

In the application below, we impose that $\omega_f + \omega_b = 0.99$ in equation (12). On one hand, this restriction provides a reasonable identifying assumption (see Mavroeidis, 2001). On the other hand, the value of this parameter is consistent with the literature on the Phillips curve.[9]

In order to investigate the consequence of mis-specification, we also consider a hybrid Phillips curve, in which additional lags of inflation are incorporated:

$$\pi_t = \omega_f E_t\pi_{t+1} + \sum_{i=1}^{3}\omega_b^i\pi_{t-i} + \beta Z_t + \varepsilon_t. \tag{13}$$

Such a model has been considered for instance by Rudebusch (2002), Mavroeidis (2001), or Galí, Gertler, and Lopez-Salido (2001). This model generalizes the model proposed by Galí and Gertler (1999). It allows firms to adopt a backward-looking rule of thumb that does not consider the one-lag inflation only, but also several lags. Note that a related approach was adopted by Fuhrer and Moore (1995a), Fuhrer (1997), Coenen and Wieland (2000), or Roberts (2001), who introduced, in addition, leads of inflation.

5.2 Empirical results

Data. The series used are displayed in Figure 1. Our sample period for estimation is 1960:I-2000:IV. Inflation is defined as the annualized quarterly percent change in the implicit GDP deflator. Output is real GDP. From a theoretical standpoint, potential output is the level that would prevail under fully flexible prices. Since estimating structural measure of potential output is beyond the scope of this paper, we concentrate on the output-gap measure computed, following Fuhrer, as the deviation of GDP from a trend with a break

[9] Indeed, in a model such as (12), the sum of the backward and forward-looking terms lies theoretically between δ and 1 and should hence be very close to 1 for any plausible value of δ (on quarterly data, δ is typically set to 0.99 in calibrated models).

in slope in 1973:IV.[10] Real marginal cost is computed, as in Galí and Gertler, using deviation of the (log) real ULC in the Non-Farm Business sector from its average value. This series is a relevant proxy for real marginal cost under the assumption of a Cobb-Douglas technology.

Estimation procedures. GMM and ML estimation procedures have been described above. We consider small variations of these procedures to take into account the standards adopted in the empirical literature. First, as previously suggested by Rudebusch (2002) or Mavroeidis (2001), several lags of inflation plays a significant role in the hybrid Phillips curve. We found that the second and third lags are necessary to fit the data correctly, while the fourth one is borderline. Second, the dynamics of the forcing variable is modelled with three lags of the inflation and the forcing variable itself. In the two cases, it was found to be sufficient to capture the dynamics of the series (see, e.g., Kurmann, 2002, for the real ULC). To be consistent with existing GMM estimates, we select instruments only dated $t-1$ or earlier, so that the instrument set contains $\{\pi_{t-1}, Z_{t-1}, ..., \pi_{t-L}, Z_{t-L}\}$, with $L = 1$ or 3. Finally, we considered different bandwidth values for the computation of the Newey and West (1987) covariance matrix. The optimal bandwidth is known to be equal to 1 (see Mavroeidis, 2001), but we also estimated the model with 12 lags, as in Galí and Gertler (1999). We found that our estimates were not altered by this change. Therefore, we report results for one lag only.

The following model is thus estimated:

$$\pi_t = \omega_f E_t \pi_{t+1} + \sum_{i=1}^{3} \omega_b^i \pi_{t-i} + \beta Z_t + \varepsilon_t \qquad (14)$$

$$Z_t = \mu_z + \sum_{i=1}^{3} \rho_i Z_{t-i} + \sum_{i=1}^{3} \theta_i \pi_{t-i} + u_t, \qquad (15)$$

with the restricted model being obtained assuming $\omega_b^2 = \omega_b^3 = 0$. Innovations are allowed to be cross-correlated. The two equations are estimated simultaneously, using the AIM procedure developed by Anderson and Moore (1985) to compute the reduced form and the log-likelihood function.

Empirical results are reported in Table 7. Panel A contains estimates of the model with real ULC, while Panel B refers to the model with output gap. In each panel, we report estimates obtained with GMM1 (one lag of the instrument set), GMM2 (three lags), ML with one single lag of inflation in the Phillips curve, and ML with three lags of inflation. The last estimates, which can be viewed as the correctly-specified model, are used below as the DGP of our Monte-Carlo simulations. Parameters of the dynamics of the forcing variable, not reported to save space, are available upon request.

It is worth noticing that, in spite of some differences in sample period and instrument set, results in Table 7 are broadly in accordance with the existing estimates both by Fuhrer and by Galí and Gertler.

[10] As an alternative to the segmented trend series, we have also used the structural potential output series computed by the Congressional Budget Office (CBO). The CBO methodology is described in Arnold (2001). Using the corresponding output gap series does not affect significantly the results (results are available upon request).

Replicating the gap between GMM and ML estimates. We begin with the model with real ULC as forcing variable (Panel A). The GMM estimates are very close to those reported by Galí and Gertler (1999) using similar method, specification, and sample. Whatever the instrument set, these estimates point to a dominant forward-looking component ($\omega_f = 0.65$). ML estimation of the same Phillips curve suggests a more balanced picture, since we obtain $\omega_f = 0.5$. Estimating a Phillips curve with three own lags indicates that the third lag is statistically significant. Yet, it does not alter other parameters as compared to the ML estimate of the one-lag model. In all estimates, the real-ULC parameter is found to be positive, yet not always significant. Note, however, that adding lags in the Newey-West covariance matrix would yield smaller standard deviation.

We turn now to the model with output gap (Panel B). As in Fuhrer (1997), we find that ML estimation provides a significant impact of output gap (with the expected positive sign) and that the backward-looking component is dominant ($\omega_f = 0.45$). It is somewhat lower than the parameter obtained by Fuhrer, since his estimates of ω_f range from 0.80 to 0.98. Note that Fuhrer does not reject the null hypothesis that the forward-looking component is "unimportant" (i.e., $\omega_f = 1$), while this restriction is rejected in our sample. This evidence contrasts quite sharply with estimates performed with GMM which all point to a non-significant parameter of output gap and to a dominant forward-looking component ($\omega_f = 0.558$ and 0.677, respectively).

Interestingly, the model with three lags of inflation suggests an even smaller weight of the forward-looking component, since ω_f is found to be as low as 0.396. As in the model with real ULC, the third lag is found to be significant.

To sum up our empirical evidence, we obtain that, with our data and our simple framework, the gap between GMM and ML estimates of ω_f is as high as $0.15 - 0.2$. This gap is present irrespective of the forcing variable that is chosen. In addition, the estimation method is found to affect the sign of the forcing-variable parameter estimate. To be consistent with our theoretical exposition above, we also estimated the two models with GMM while including both real ULC and output gap in the instrument set. This is a common practice, in the following of Galí and Gertler (1999) and Roberts (2001), and reflecting the practice of an econometrician uncertain about the true DGP. We found that the estimation of the Phillips curve with real ULC was not altered at all. As far as the output-gap model is concerned, we obtained an even larger forward-looking component ($\omega_f = 0.7$), yet a negative output-gap parameter. (results are not reported to save space.)

5.3 Filling the gap

In this section, we investigate the source of the discrepancy between the two estimation methods with some Monte-Carlo simulations. On the basis of the previous experiments, we assume that the true DGP is likely to be close to the estimated obtained with three lags of inflation. Indeed, only a small value of ω_f with an omitted dynamic can reconcile the empirical evidence with our experiments. In this case, indeed, we would obtain a significant bias of GMM as well as ML estimators, but the GMM estimator would be biased toward 1, while the ML estimator would be biased toward 0.5 only.

In order to simplify the exposition, we consider a model in which the second lag (rather than the third one) of inflation is relevant, while omitted in the estimated Phillips curve. Also, we specify a model for the forcing variable that contains one lag only. Then, the DGP is the following: $\omega_f = 0.4$, $\omega_b^1 = 0.5$, $\omega_b^2 = 0.1$, $\beta = 0.1$, $\rho = 0.8$, $\theta = -0.05$.

Values for ρ and θ correspond to those obtained with the output-gap model. Last, we set $\sigma_u^2 = \sigma_\varepsilon^2 = 1$. **Table 8** reports the results of these simulations. Panel A is devoted to the Plim of the estimators. They are broadly consistent with estimates reported in Table 7 Panel B. An important exception is the excessive Plim of ω_f obtained with GMM2 (0.746, while the empirical estimate is 0.677). Panel B reports results on the finite-sample distribution of the estimators (computed with 1000 samples of $T = 150$ observations). The most interesting result is that the GMM2 bias of ω_f reported in Panel A decreases very significantly (to 0.65), so that it is now very close to the empirical estimate. This result reflects the fact that the finite-sample bias partly offsets the (asymptotic) mis-specification bias. The former bias is related to the weak instrument relevance, which occurs in this instance because GMM2 includes too many instruments, some of which are irrelevant. Therefore, the estimator is biased toward the Plim of the OLS estimator, which is known to be close to 0.5. On the whole, the finite-sample performance of GMM2 appears to benefit from the combination of two biases of opposite signs.

6 Conclusion

This paper has analyzed the properties of GMM and ML estimators in hybrid models. Our motivation was the gap between the large degree of forward-looking behavior typically found when implementing GMM and the low degree of forward-lookingness obtained by ML. Our findings can be summarized as follows. First, finite-sample biases are not able to fill the gap between empirical estimates. The GMM bias is small unless a large number of irrelevant variables are used. Furthermore, the bias is toward the Plim of the OLS which is close to 0.5 in most cases, and the estimator may take on a lower value than the ML in small samples.

Second, plausible mis-specifications can produce substantial differences between these two estimators. In particular, in case of measurement error, GMM can be moderately biased toward 1. A relevant condition for this feature to appear is that the relevant variable is omitted from the estimated equation but included in the instrument set. Analytical results establish that, in a simple model without feedback, asymptotic ML and GMM biases are in opposite directions from the true value.

In case of omitted dynamic, the GMM estimators are likely to be severely biased toward implausibly large values in case of a large forward-looking components. In general, however, biases of GMM and ML point to the same direction. While this property does not carry on to more elaborate models with feedback, we still find that, in case of mis-specification, GMM is more widely biased than ML in a way that is likely to fill the gap between these estimators.

The results in the present paper are used to rationalize differences in estimation of the Phillips curve found in the literature. Small mis-specification (such as omitting one relevant lag with a modest parameter value in the inflation dynamics) turns out to imply substantial over-estimation of the degree of forward-lookingness found by the GMM estimator.

7 Appendices

7.1 Appendix 1: The case with measurement error

7.1.1 Description of the problem

The DGP is assumed to be described by the following equations:

$$\begin{aligned} Y_t &= \omega_f E_t Y_{t+1} + (1-\omega_f) Y_{t-1} + \beta Z_t + \varepsilon_t && (16) \\ Z_t &= \rho Z_{t-1} + u_t && (17) \\ X_t &= a Z_t + e_t && (18) \end{aligned}$$

where Y_t is the variable of interest, Z_t is the true forcing variable, X_t is the proxy used by the econometrician. All innovations, ε_t, u_t, and e_t are contemporaneously and serially uncorrelated. We denote $\sigma_\varepsilon^2 = E\left(\varepsilon_t^2\right)$, $\sigma_u^2 = E\left(u_t^2\right)$, and $\sigma_e^2 = E\left(e_t^2\right)$. We also denote $\tau = a^2\sigma_Z^2/\left(a^2\sigma_Z^2 + \sigma_e^2\right)$ the fraction of the variance of X_t explained by Z_t. We impose that coefficients on lead and lag sum to one in order to insure identification of the model parameters (see Mavroeidis, 2001, for a discussion).

Denoting φ_1 and φ_2 the roots of the characteristic polynomial in equation (16), we have

$$(1-\varphi_1 L)(1-\varphi_2 F) Y_t = \frac{\beta}{\omega_f} Z_t + \frac{\varepsilon_t}{\omega_f}.$$

Using the dynamic of Z_t given by equation (17) gives

$$(1-\varphi_1 L) Y_t = \frac{\beta}{\omega_f} \sum_{i=0}^{\infty} \rho^i Z_t + \frac{\varepsilon_t}{\omega_f},$$

so that the reduced form of the DGP is:

$$Y_t = \varphi_1 Y_{t-1} + \theta Z_t + \tilde{\varepsilon}_t \tag{19}$$

where $\varphi_1 = (1-\omega_f)/\omega_f$, $\varphi_2 = 1$, $\theta = \beta/\left(\omega_f(1-\rho)\right)$, and $\tilde{\varepsilon}_t = \varepsilon_t/\omega_f$. Stationarity of the model requires that $|\varphi_1| < 1$ (or, equivalently, $\omega_f > 0.5$) and $|\rho| < 1$.

The estimated mis-specified model is:

$$\begin{aligned} Y_t &= \alpha_f E_t Y_{t+1} + (1-\alpha_f) Y_{t-1} + b X_t + \upsilon_t && (20) \\ X_t &= \psi X_{t-1} + w_t. && (21) \end{aligned}$$

The econometrician estimates α_f as the degree of forward-lookingness. It is a presumably a biased estimator of the parameter ω_f. Note that there is no mi-specification in the limiting case where $\sigma_e^2 = 0$, i.e. $\tau = 1$.

We now consider the GMM and ML estimators of parameters α_f and b. Structural parameters are $\{\omega_f, \beta, \rho, a, \sigma_\varepsilon^2, \sigma_u^2, \sigma_e^2\}$.

7.1.2 Some formulae on cross-moments

We now report several cross-moment coefficients implied by the DGP. They are useful to compute the moment conditions. Moments of Z_t and cross-moments with Y_t are:

$$EZ_t^2 = \frac{\sigma_u^2}{1-\rho^2} = \sigma_Z^2$$

$$
\begin{aligned}
EZ_tZ_{t-1} &= \rho\sigma_Z^2 \\
EZ_tY_t &= \frac{\theta\sigma_Z^2}{1-\varphi_1\rho} = \Phi_0 \\
EZ_tY_{t-1} &= \rho\Phi_0 \\
EY_tZ_{t-1} &= \varphi_1\Phi_0 + \theta\rho\sigma_Z^2.
\end{aligned}
$$

Moments involving only Y_t are:

$$
\begin{aligned}
EY_t^2 &= \frac{\sigma_\varepsilon^2}{1-\varphi_1^2} + \frac{1+\varphi_1\rho}{1-\varphi_1^2}\theta\Phi_0 = \Gamma_0 \\
EY_tY_{t-1} &= \varphi_1\Gamma_0 + \theta\rho\Phi_0 \\
EY_tY_{t-2} &= \varphi_1^2\Gamma_0 + \theta\rho\left(\varphi_1+\rho\right)\Phi_0.
\end{aligned}
$$

Finally, moments of X_t and cross-moments with Y_t are:

$$
\begin{aligned}
EX_t^2 &= a^2\sigma_Z^2 + \sigma_e^2 \\
EX_tX_{t-1} &= \rho a^2\sigma_Z^2 \\
EX_tY_{t-i} &= aEZ_tY_{t-i} \qquad \forall i.
\end{aligned}
$$

7.1.3 OLS in the correctly-specified case

Some authors (Nelson and Startz, 1990, Staiger and Stock, 1997, and Woglow, 2001) have highlighted that including several variables in the instrument set is likely to bias the GMM estimator within small samples. More precisely, when the number of instruments increases with the sample size, the estimator has the same probability limit as the OLS estimator, which is known to be biased in such a situation. The probability limit of the OLS estimator is given by

$$
\begin{aligned}
\alpha_{OLS} &= \frac{1}{2}\left(\frac{\sigma_\varepsilon^2 + \Lambda^2\left[1-\varphi_1\right]}{\sigma_\varepsilon^2 + \Lambda^2\left[\frac{1}{2}\left(1-\varphi_1^2\right) + \left(1-\varphi_1\rho\right)^2\right]}\right) \\
b_{OLS} &= \left(\frac{\beta}{a}\right)\frac{1+\varphi_1}{1-\rho}\left(\frac{\sigma_\varepsilon^2 + \Lambda^2\left[\frac{2(1-\rho)-\varphi_1(1-\rho^2)}{1-\varphi_1\rho}\right]}{\sigma_\varepsilon^2 + \Lambda^2\left[\frac{1}{2}\left(1-\varphi_1^2\right) + \left(1-\varphi_1\rho\right)^2\right]}\right),
\end{aligned}
$$

where $\Lambda = \beta\sigma_u\left(1+\varphi_1\right)/\left[\left(1-\varphi_1\rho\right)\left(1-\rho\right)\right]$. For a wide range of structural parameters, the OLS estimator of α_f is biased toward 0.5, although the Plim is not exactly 0.5. More generally, the Plim above indicates that $0 \leq \alpha_{OLS} \leq 0.5$. Such a result has been obtained numerically in several Monte-Carlo studies, but not yet analytically. See, for instance, Lindé (2001), Mavroeidis (2001), and Fuhrer and Rudebusch (2002).

7.1.4 The GMM estimators

Two parameters α_f and b are estimated, so that at least two instruments are needed. We consider two cases: estimator GMM1 is based on the instrument set $\{Y_{t-1}, X_t\}$, while estimator GMM2 is based on the instrument set $\{Y_{t-1}, X_t, Z_t\}$. In the second case, the instruments include the actual forcing variable and there is over-identification.

GMM1. GMM1 relies on the following moment conditions:

$$\begin{aligned} E\left[Y_{t-1}(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f)\, Y_{t-1} - bX_t)\right] &= 0 \\ E\left[X_t(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f)\, Y_{t-1} - bX_t)\right] &= 0. \end{aligned}$$

Since the model is just-identified, the probability limits of the estimators α_f and b can be obtained by solving these moment conditions, so that

$$\alpha_{GMM1} = \frac{EY_t^2 - EY_tY_{t-1} + \left(EX_tY_t/EX_t^2\right)\left(EX_tY_t - EX_tY_{t-1}\right)}{EY_t^2 - EY_tY_{t-2} + \left(EX_tY_t/EX_t^2\right)\left(EX_tY_{t-1} - EX_{t-1}Y_t\right)}.$$

After some cumbersome computations, we obtain:

$$\alpha_{GMM1} = \omega_f \left(\frac{\sigma^2 + \Lambda^2\left[1 - (1-\tau)\frac{\rho(1+\varphi_1)}{1+\rho}\right]}{\sigma^2 + \Lambda^2\left[1 - (1-\tau)\rho\varphi_1\right]} \right).$$

The first term is the asymptotic value of α_f when the model is correctly specified. We note that the bias on α_{GMM1} has the same sign as $\rho\varphi_1(1+\rho) - \rho(1+\varphi_1) = -\rho(1-\rho\varphi_1)$. Since this expression is always negative, we conclude that α_{GMM1} is biased toward zero.

The estimator of b is obtained by replacing α_f by α_{GMM1} in the expression

$$E\left[X_t\left(Y_t - \alpha_f E_t Y_{t+1} - (1-\alpha_f)\, Y_{t-1} - bX_t\right)\right] = 0,$$

yielding

$$b_{GMM1} = \left(\frac{\beta}{a}\tau\right)\left(\frac{\sigma^2 + \Lambda^2}{\sigma^2 + \Lambda^2\left[1 - (1-\tau)\rho\varphi_1\right]}\right).$$

Note that this estimator should not be compared to β directly, since X_t may be a good proxy for Z_t yet having a parameter a different from 1. Therefore, assessing the bias for b involves comparing the estimator to β/a. In addition, since X_t is only a proxy of Z_t, b can be, at best, an estimator of $(\beta\tau/a)$. Therefore, two sources of bias operate: The first bias comes from τ, which measures whether X_t is a good proxy for Z_t. The second bias comes from the second term between brackets, which measures the extent of the mis-specification bias. While the first component biases b_{GMM1} toward zero, the second bias is positive. Assuming $\tau = 0$ yields $b_{GMM1} = 0$, whereas assuming $\tau = 1$ yields $b_{GMM1} = \beta/a$. We verify that the whole bias is negative.

GMM2. Since estimator GMM2 is over-identified, it proves convenient to view it as a two-step estimator. First, we regress Y_{t+1} on the instrument set to build the fitted value $\widehat{Y}_{t+1}$. Since Z_t is the true forcing variable, this yields $\hat{Y}_{t+1} = \varphi_1^2 Y_{t-1} + \theta(\varphi_1 + \rho)\, Z_t$. Then, the second-stage regression has the two following moment conditions:

$$\begin{aligned} &E\left[\left(\hat{Y}_{t+1} - Y_{t-1}\right)\left(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f)\, Y_{t-1} - bX_t\right)\right] = 0 \\ &E\left[X_t\left(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f)\, Y_{t-1} - bX_t\right)\right] = 0. \end{aligned}$$

After some computations, we obtain:

$$\alpha_{GMM2} = \omega_f \left(\frac{\sigma_\varepsilon^2 + \Lambda^2\left[1 + (1-\tau)\frac{(1-\rho)\varphi_1}{1-\varphi_1}\right]}{\sigma_\varepsilon^2 + \Lambda^2\left[1 + (1-\tau)\frac{(1-\rho^2)\varphi_1^2}{(1-\varphi_1^2)}\right]} \right).$$

The bias on α_{GMM2} has the same sign as $(1-\rho)\,\varphi_1\,(1-\varphi_1^2) - (1-\rho^2)\,\varphi_1^2\,(1-\varphi_1) = \varphi_1\,(1-\rho)\,(1-\rho\varphi_1)$. Since this expression is always positive when $\rho\varphi_1 < 1$, we conclude that α_{GMM2} is biased toward one.

The estimator of b is obtained by replacing α_f by α_{GMM2} in the expression

$$E\left[X_t\left(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f)\,Y_{t-1} - bX_t\right)\right] = 0.$$

This yields

$$b_{GMM2} = \left(\frac{\beta}{a}\tau\right)\left(\frac{\sigma_\varepsilon^2 + \Lambda^2}{\sigma_\varepsilon^2 + \Lambda^2\left[1 + (1-\tau)\frac{(1-\rho^2)\varphi_1^2}{(1-\varphi_1^2)}\right]}\right).$$

In this case, both components point to a bias toward zero. Therefore, the whole bias is negative. We notice that the differences $(\alpha_{GMM2} - \alpha_{GMM1})$ and $(b_{GMM1} - b_{GMM2})$ are always positive.

7.1.5 The ML estimator

This estimator is based on the reduced form of the postulated system (20)–(21):

$$Y_t = \varphi Y_{t-1} + \mu X_{t-1} + \tilde{\varepsilon}_t.$$

Parameters φ and μ can be estimated by OLS:

$$\begin{pmatrix} \varphi_{ML} \\ \mu_{ML} \end{pmatrix} = \begin{pmatrix} \frac{EX_t^2 EY_tY_{t-1} - EX_tY_{t-1}EX_tY_t}{EX_t^2EY_t^2 - (EX_tY_{t-1})^2} \\ \frac{EY_t^2 EY_tX_t - EX_tY_{t-1}EY_tY_{t-1}}{EX_t^2EY_t^2 - (EX_tY_{t-1})^2} \end{pmatrix}.$$

After some computations, we obtain:

$$\varphi_{ML} = \varphi_1\left(\frac{\sigma_\varepsilon^2 + \Lambda^2\left[(1-\rho^2) + \frac{\rho}{\varphi_1}(1-\tau)(1-\varphi_1^2)\right]}{\sigma_\varepsilon^2 + \Lambda^2\left[(1-\rho^2) + \rho^2(1-\tau)(1-\varphi_1^2)\right]}\right)$$

and

$$\mu_{ML} = \left(\frac{\beta}{a}\tau\right)\frac{1+\varphi_1}{1-\rho}\left(\frac{\sigma_\varepsilon^2 + \Lambda^2}{\sigma_\varepsilon^2 + \Lambda^2\left[1 + (1-\tau)\frac{(1-\varphi_1^2)\rho^2}{(1-\rho^2)}\right]}\right).$$

Then, the "structural" parameters $\alpha_{_{ML}}$ and b_{ML} are given by the condition $\alpha_{ML} = 1/\left(1+\varphi_{ML}\right)$ and $b_{ML} = \mu_{ML}\alpha_{_{ML}}\left(1-\psi\right)$, so that:

$$\alpha_{ML} = \omega_f\left(\frac{\sigma^2 + \Lambda^2\left[1 + (1-\tau)\frac{(1-\varphi_1^2)\rho^2}{1-\rho^2}\right]}{\sigma^2 + \Lambda^2\left[1 + (1-\tau)\frac{(1-\varphi_1)\rho}{1-\rho}\right]}\right)$$

and

$$b_{ML} = \left(\frac{\beta}{a}\tau\right)\frac{1-\psi}{1-\rho}\left(\frac{\sigma^2 + \Lambda^2}{\sigma^2 + \Lambda^2\left[1 + (1-\tau)\,\frac{(1-\varphi_1)\rho}{1-\rho}\right]}\right).$$

Therefore, the bias on α_{ML} has the same sign as $(1-\varphi_1^2)\rho^2 - (1-\varphi_1)\rho\,(1+\rho) = -(1-\varphi_1)\rho\,(1-\varphi_1\rho)$ which is negative when $\varphi_1\rho < 1$. We obtain three sources of bias for b_{ML}: The first bias comes from τ, which measures whether X_t is a good proxy for Z_t. The second bias comes from the expression $(1-\psi)/(1-\rho)$, with $\psi = \rho\tau$, which measures the mis-measurement in the serial correlation of the forcing variable. The last bias comes from the second term between brackets, which measures the mis-specification bias. All components are smaller than one, except the second term. Therefore, the whole bias is of either sign. When $\tau = 0$, we obtain $b_{ML} = 0$, whereas when $\tau = 1$, we obtain $b_{ML} = \left(\frac{\beta}{a}\right)/(1-\rho) > \left(\frac{\beta}{a}\right)$.

We have $\alpha_{ML} > \alpha_{GMM1}$ when $\rho < \sigma^2\varphi_1/\left(\sigma^2 - \left(1-\varphi_1^2\right)(1-\tau)\right)$. This is likely to be the case when ω_f and τ are small. We also notice that the difference $(b_{ML} - b_{GMM1})$ is always positive.

7.2 Appendix 2: The case with omitted dynamics

7.2.1 Description of the problem

The DGP is assumed to be described by the following equations:

$$Y_t = \omega_f E_t Y_{t+1} + \omega_b^1 Y_{t-1} + \left(1-\omega_f-\omega_b^1\right) Y_{t-2} + \beta Z_t + \varepsilon_t \tag{22}$$

$$Z_t = \rho Z_{t-1} + u_t \tag{23}$$

where Y_t is the variable of interest, Z_t is the true forcing variable. Innovations, ε_t and u_t, are contemporaneously and serially uncorrelated. We denote $E\left(\varepsilon_t^2\right) = \sigma_\varepsilon^2$, and $E\left(u_t^2\right) = \sigma_u^2$. We impose that coefficients on lead and lags sum to one in order to insure identification of the model parameters.

Denoting δ_1, δ_2, and δ_3 the roots of the characteristic polynomial in equation (22), we have

$$(1-\delta_1 L)(1-\delta_2 L)(1-\delta_2 F) Y_t = \frac{\beta}{\omega_f} Z_t + \frac{1}{\omega_f}\varepsilon_t.$$

The reduced form of this DGP is given by the expression

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \theta Z_t + \tilde{\varepsilon}_t \tag{24}$$

where $\varphi_1 = \delta_1 + \delta_2$ and $\varphi_2 = -\delta_1\delta_2$. Reduced-form parameters and DGP parameters are related by the following relationships: $\varphi_1 = (1-\omega_f)/\omega_f$, $\varphi_2 = \left(1-\omega_f-\omega_b^1\right)/\omega_f$, $\varphi_3 = 1$, $\theta = \beta/\left(\omega_f(1-\rho)\right)$, and $\tilde{\varepsilon}_t = \varepsilon_t/\omega_f$.

Stationarity conditions of an AR(2) process such as equation (24) are known to be: $1-\varphi_1-\varphi_2 < 1$, $1+\varphi_1-\varphi_2 > 0$, and $\varphi_2 > -1$. These conditions are equivalent to the conditions $\left(2-3\omega_f-\omega_b^1\right) < 0$ and $-\omega_f < \omega_b^1 < 1$. **Figure 2** displays, in the plane $\left\{\omega_f, \omega_b^1\right\}$, the area where the hybrid equation is stationary. It is defined by the constraints $\left(2-3\omega_f-\omega_b^1\right) < 0$ and $-\omega_f < \omega_b^1 < 1$. The segment $\{(1,-1),(1/3,1)\}$ corresponds to the case of non-stationarity with $\varphi_1+\varphi_2 = 1$. Along this segment, ω_b^2 decreases from $\omega_b^2 = 1$ to $-1/3$. Notice that choosing $\omega_b^1 = 1$ does not imply any restriction upon persistence of Y_t, since ω_f may itself vary. The segment $\{(1/3,1),(1,1)\}$ is consistent with $\varphi_1+\varphi_2$ ranging from 1 to -1 (ω_b^2 ranges from $-1/3$ to -1). In the following Monte-Carlo experiments, we will focus essentially to the area $\{(1/3,1),(1,1),(1,-1)\}$ assuming that $\omega_f \leq 1$. This assumption is not necessary from a statistical viewpoint but economic interpretation generally requires $\omega_f \in [0,1]$. Note, in addition, that assuming that ω_f, ω_b^1, and ω_b^2 are all between $[0,1]$ would imply to restrict to the area $\left(\omega_f, \omega_b^1\right) = \{(1/2,1/2),(1,0),(1,-1)\}$. This area also corresponds to the cases where $\varphi_1, \varphi_2 \geq 0$.

The estimated mis-specified model is:

$$Y_t = \alpha_f E_t Y_{t+1} + (1-\alpha_f) Y_{t-1} + bZ + v_t. \tag{25}$$

The econometrician estimates α_f as the degree of forward-lookingness. It is presumably a biased estimator of parameter ω_f. Note that there is no mis-specification in the limiting case where $\varphi_2 = 0$, i.e. $\omega_f + \omega_b^1 = 1$. We now consider the GMM and ML estimators of parameters α_f and b. Structural parameters are $\{\omega_f, \omega_b^1, \beta, \rho, a, \sigma_\varepsilon^2, \sigma_u^2\}$.

7.2.2 Some formulas on cross-moments

We now report several cross-moment coefficients implied by the DGP. Moments of Z_t are the same as in the measurement-error case, while cross-moments with Y_t are:

$$\begin{aligned} EZ_tY_t &= \frac{\theta\sigma_Z^2}{1-\varphi_1\rho-\varphi_2\rho^2} = \Phi_0 \\ EZ_tY_{t-1} &= \rho\Phi_0 \\ EY_tZ_{t-1} &= (\varphi_1+\varphi_2\rho)\,\Phi_0 + \theta\rho\sigma_Z^2. \end{aligned}$$

Moments involving Y_t only are:

$$\begin{aligned} EY_t^2 &= \frac{(1-\varphi_2)\,\sigma_\varepsilon^2 + \left[\varphi_1\rho\,(1+\varphi_2) + (1-\varphi_2)\,(1+\varphi_2\rho^2)\right]\theta\Phi_0}{(1+\varphi_2)\,(1-\varphi_1-\varphi_2)\,(1+\varphi_1-\varphi_2)} = \Gamma_0 \\ EY_tY_{t-1} &= \frac{\varphi_1}{1-\varphi_2}\Gamma_0 + \frac{\theta\rho}{1-\varphi_2}\Phi_0 \\ EY_tY_{t-2} &= \frac{\sigma_\varepsilon^2}{1+\varphi_2} + \theta\frac{1-\rho^2}{1+\varphi_2}\Phi_0. \end{aligned}$$

7.2.3 The GMM estimators

Two parameters α_f and b are estimated, so that at least two instruments are needed. We consider two cases: Estimator GMM1 is based on the instrument set $\{Y_{t-1}, Z_t\}$, while estimator GMM2 is based on the instrument set $\{Y_{t-1}, Y_{t-2}, Z_t\}$. In the latter case, the instrument set includes a second lag of the dependent variable and there is over-identification.

GMM1. Estimator GMM1 relies on the following moment conditions:

$$\begin{aligned} E\left[Y_{t-1}(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f)\,Y_{t-1} - bZ_t)\right] = 0 \\ E\left[Z_t(Y_t - \alpha_f Y_{t+1} - (1-\alpha_f)\,Y_{t-1} - bZ_t)\right] = 0. \end{aligned}$$

Since the model is just-identified, the probability limits of the estimators of α_f and b can be obtained by the two moment conditions, so that

$$\alpha_{GMM1} = \frac{EY_t^2 - EY_tY_{t-1} + \left(EZ_tY_t/EZ_t^2\right)(EZ_tY_t - EZ_tY_{t-1})}{EY_t^2 - EY_tY_{t-2} + \left(EZ_tY_t/EZ_t^2\right)(EZ_tY_{t-1} - EZ_{t-1}Y_t)}.$$

After some cumbersome computations, we obtain:

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

where $\tilde{\Lambda} = \beta\sigma_u (1+\varphi_1) / \left[\left(1-\varphi_1\rho-\varphi_2\rho^2\right)(1-\rho)\right]$. The first term of α_{GMM1} is the asymptotic value of α_f when $\rho = 0$.

The estimator of b is obtained by replacing α_f by α_{GMM1} in the expression

$$E\left[Z_t \left(Y_t - \alpha_f E_t Y_{t+1} - (1-\alpha_f) Y_{t-1} - bZ_t\right)\right] = 0,$$

yielding

$$\begin{aligned} b_{GMM1} &= \beta \left(\frac{(1+\varphi_1)\left(1-\varphi_1\rho-\varphi_2\rho^2-\varphi_2(1+\rho)\right)}{(1+\varphi_1-\varphi_2)\left(1-\varphi_1\rho-\varphi_2\rho^2\right)} \right) \\ &\times \left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2 \left[1+\varphi_2\rho(\varphi_1+\varphi_2\rho)\frac{(1-\varphi_1\rho-\varphi_2\rho^2)+(1+\rho)}{(1-\varphi_1\rho-\varphi_2\rho^2)+(1+\rho)\varphi_2}\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho(\varphi_1+\varphi_2\rho)\right]} \right). \end{aligned}$$

GMM2. As in the case of measurement error, since estimator GMM2 is over-identified, it is estimated in two steps: First, we regress Y_{t+1} on the instrument set to build the fitted value $\widehat{Y}_{t+1}$. Then, we estimate, by OLS:

$$Y_t - Y_{t-1} = \alpha_f \left(\widehat{Y}_{t+1} - Y_{t-1}\right) + bZ_t + \xi_t.$$

Since Z_t is the true forcing variable, this yields $\hat{Y}_{t+1} = \varphi_1^2 Y_{t-1} + \theta(\varphi_1+\rho) Z_t$. Then, the second-stage regression has the two following moment conditions:

$$\begin{aligned} &E\left[\left(\hat{Y}_{t+1} - Y_{t-1}\right)\left(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f) Y_{t-1} - bX_t\right)\right] = 0 \\ &E\left[X_t\left(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f) Y_{t-1} - bX_t\right)\right] = 0. \end{aligned}$$

After some computations, we obtain:

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

The estimator of b is obtained by replacing α_f by α_{GMM2} in the expression

$$E\left[X_t\left(Y_t - \alpha_f \hat{Y}_{t+1} - (1-\alpha_f) Y_{t-1} - bX_t\right)\right] = 0,$$

yielding

$$\begin{aligned} b_{GMM2} &= \beta \left(\frac{(1+\varphi_1)\left(1-\varphi_1-\varphi_2-\rho(1-\varphi_1-\varphi_1\varphi_2)(\varphi_1+\varphi_2+\varphi_2\rho)\right)}{\left(1-\varphi_1\rho-\varphi_2\rho^2\right)(1-\varphi_1-\varphi_1\varphi_2-\varphi_2)} \right) \\ &\times \left(\frac{\sigma_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{2\varphi_1+\varphi_2\rho-\varphi_2\rho-(\varphi_1+\varphi_2\rho)(-1+\varphi_1+\varphi_2-\rho(1-\varphi_1+\varphi_2\rho))}{1-\varphi_1-\varphi_2-\rho(1-\varphi_1-\varphi_1\varphi_2)(\varphi_1+\varphi_2+\varphi_2\rho)}\right]}{\sigma_\varepsilon^2+\tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{2\varphi_1+\varphi_2\rho-\varphi_2\rho}{1-\varphi_1-\varphi_1\varphi_2-\varphi_2}\right]} \right). \end{aligned}$$

7.2.4 The ML estimator

This estimator is based on the reduced form of the postulated system (22)–(23):

$$Y_t = \varphi Y_{t-1} + \mu Z_{t-1} + \tilde{\varepsilon}_t.$$

Parameters φ and μ can be estimated by OLS:

$$\begin{pmatrix} \varphi_{ML} \\ \mu_{ML} \end{pmatrix} = \begin{pmatrix} \frac{EZ_t^2 EY_t Y_{t-1} - EZ_t Y_{t-1} EZ_t Y_t}{EZ_t^2 EY_t^2 - (EZ_t Y_{t-1})^2} \\ \frac{EY_t^2 EY_t Z_t - EZ_t Y_{t-1} EY_t Y_{t-1}}{EZ_t^2 EY_t^2 - (EZ_t Y_{t-1})^2} \end{pmatrix}.$$

Then, the "structural" parameters α_{ML} and b_{ML} are given by the condition $\alpha_{ML} = 1/\left(1+\varphi_{ML}\right)$ and $b_{ML} = \mu_{ML}\alpha_{_{ML}}\left(1-\psi\right)$. After some computations, we obtain:

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

and

$$\begin{aligned} b_{ML} &= \beta\left(\frac{\left(1+\varphi_1\right)\left(1+\varphi_1-\varphi_2-\rho\left(1-\varphi_2\right)\right)}{\left(1-\varphi_1-\varphi_2\right)\left(1-\varphi_1\rho-\varphi_2\rho^2\right)}\right) \\ &\quad\times\left(\frac{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{(1-\rho)(2\varphi_1+\varphi_2\rho-\varphi_2\rho)+1+\varphi_1-\varphi_2+\varphi_1\varphi_2\rho}{1+\varphi_1-\varphi_2-\rho(1-\varphi_2)}\right]}{\sigma_\varepsilon^2 + \tilde{\Lambda}^2\left[1+\varphi_2\rho\frac{(2\varphi_1+\varphi_2\rho-\varphi_2\rho)+1+\varphi_1-\varphi_2+\varphi_1\varphi_2\rho}{1-\varphi_1-\varphi_2}\right]}\right). \end{aligned}$$

References

[1] Amato, J. D., and T. Laubach (2003), Rule-of-Thumb Behaviour and Monetary Policy, forthcoming in *European Economic Review*.

[2] Anderson, G. A., and G. R. Moore (1985), A Linear Algebraic Procedure for Solving Perfect Foresight Models, *Economics Letters*, 17, 247–252.

[3] Arnold, R. (2001), CBO's method for Estimating Potential Output: An Update, CBO Memorandum, August 2001.

[4] Calvo, G. (1983), Staggered Prices in a Utility-Maximizing Framework, *Journal of Monetary Economics*, 12(3), 383–398.

[5] Chadha, B., P. Masson, and M. Meredith (1992), Models of Inflation and the Cost of Disinflation *IMF Staff Paper*, 39(2), 395-431.

[6] Christiano, L. J., M. Eichenbaum, and C. Evans (2001), Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, Federal Reserve Bank of Cleveland Working Paper 01-07.

[7] Coenen, G., and V. Wieland (2000), A Small Estimated Euro Area Model with Rational Expectations and Nominal Rigidities, ECB Working Paper 30.

[8] Fuhrer, J. C. (1997), The (Un)importance of Forward Looking Behavior in Price Specifications, *Journal of Money, Credit, and Banking*, 29(3), 338–350.

[9] Fuhrer, J. C. (2000), Habit Formation in Consumption and Its Implications for Monetary-Policy Models, *American Economic Review*, 90(3), 367–390.

[10] Fuhrer, J. C., and G. R. Moore (1995a), Monetary Policy Trade-Off and the Correlation Between Nominal Interest Rates and Real Output, *American Economic Review*, 85(1), 219–239.

[11] Fuhrer, J. C., and G. R. Moore (1995b), Inflation Persistence, *Quarterly Journal of Economics*, 110(1), 127–160.

[12] Fuhrer, J. C., G. R. Moore, and S. D. Schuh (1995), Estimating the Linear Quadratic Inventory Model. Maximum Likelihood versus Generalized Method of Moments, *Journal of Monetary Economics*, 35(1), 115–157.

[13] Fuhrer, J. C., and Rudebusch, G. D. (2002) Estimating the Euler equation for output, Federal Reserve Bank of Boston Working Paper 02-3.

[14] Galí, J., and M. Gertler (1999), Inflation Dynamics: A Structural Econometric Analysis, *Journal of Monetary Economics*, 44(2), 195–222.

[15] Galí, J., M. Gertler, and J. D. López-Salido (2001), European Inflation Dynamics, *European Economic Review*, 45(7), 1237–1270.

[16] Guerrieri, L. (2001), Inflation Dynamics, Board of Governors of the Federal Reserve System, International Finance Discussion Paper 715.

[17] Hall, A. R., G. D. Rudebusch, and D. W. Wilcox (1996), Judging Instrument Relevance in Instrumental Variables Estimation, *International Economic Review*, 37(2), 283–298.

[18] Klein, P. (2000), Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model, *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

[19] Kurmann, A. (2002), Can New Keynesian Pricing Explain U.S. Inflation Dynamics?, University of Virginia, Working Paper.

[20] Lindé, J. (2001), Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach, Sveriges Riksbank Working Paper Series 129.

[21] Mavroeidis, S. (2001), Identification and Mis-Specification Issues in Forward-Looking Monetary Models, Nuffield College, Working Paper.

[22] Nelson, C. R., and R. Startz (1990), Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator, *Econometrica*, 58(4), 967–976.

[23] Newey, C. R., and K. D. West (1987), A Simple, Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 55(3), 703–708.

[24] Oliner, S. D., G. D. Rudebusch, and D. Sichel (1996), The Lucas Critique Revisited: Assessing the Stability of Empirical Euler Equations for Investment, *Journal of Econometrics*, 70(1), 291–316.

[25] Otrok, C, B. Ravikumar, and C. H. Whiteman (2002) Habit Formation: A Resolution of the Equity Premium Puzzle?, *Journal of Monetary Economics*, 49(6), 1261–1288.

[26] Roberts, J. M. (2001), How Well Does the New Keynesian Sticky-Price Model Fit the Data?, Federal Reserve Board, Finance and Economics Discussion Series, 2001-13.

[27] Rotemberg, J. (1982), Monopolistic Price Adjustment and Aggregate Output, *Review of Economic Studies*, 49(4), 517–531.

[28] Rudd, J., and K. Whelan (2001), New Tests of the New-Keynesian Phillips Curve, Finance and Economics Discussion Series, FEDS, 2001-30.

[29] Rudebusch, G. D. (2002), Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia, *Journal of Monetary Economics*, 49(6),1161–1187.

[30] Sawa, T. (1978), The Exact Moment of the Least Square Estimator for the Autoregressive Model, *Journal of Econometrics*, 8(2), 159–172.

[31] Sbordone, A. M. (2002), Prices and Unit Labor Cost: A New Test of Price Stickiness, *Journal of Monetary Economics*, 49(2), 265–292.

[32] Söderlind, P. (1999), Solution and Estimation of RE Macromodels with Optimal Policy, *European Economic Review*, 43(4-6), 813–823.

[33] Staiger, D., and J. H. Stock (1997), Instrumental Variables Regression with Weak Instruments, *Econometrica*, 65(3), 557–586.

[34] Taylor, J. B. (1980), Aggregate Dynamics and Staggered Contracts, *Journal of Political Economy*, 88(1), 1–23.

[35] Tinsley, P. A. (1993), Fitting both Data and Theories: Polynomial Adjustment Costs and Error-Correction Decision Rules, Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 93-21.

[36] Tinsley, P. A. (2000), Rational Error Correction, University of Cambridge, Working Paper.

[37] Woglom, G. (2001), More Results on the Exact Small Sample Properties of the Instrumental Variable Estimator, *Econometrica*, 69(5), 1381–1389.

Captions

Table 1: This table reports finite-sample estimators of the model with a single lag. Parameter sets are $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, and $\sigma_u = \sigma_\varepsilon = 1$. The median and the MAD of the parameter distribution are computed over $N = 2000$ simulations of size $T = 100$. Estimation methods are GMM with $L = 0$ and 7 lags of $W_t = \{Y_{t-1}, Z_t\}$ and ML. The Plim of the OLS estimator is also computed with a sample of $100,000$ observations.

Table 2: This table reports Plim of GMM and ML estimators in the case of measurement error. Parameter sets are $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, $\sigma_u = \sigma_\varepsilon = 1$, and $\tau = \{0.1; 0.5; 0.9\}$. These Plims are computed using Proposition 1.

Table 3: This table reports Plim of GMM and ML estimators in the case of omitted dynamics. Parameter sets are $\omega_f = \{0.4; 0.65; 0.9\}$, $\rho = \{0; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, $\sigma_u = \sigma_\varepsilon = 1$, and ω_b^1 and ω_b^2 chosen such that persistence $\varphi_1 + \varphi_2 = \{0; 0.5; 0.9\}$. These Plims are computed using Proposition 2.

Table 4: This table reports finite-sample estimators of the model with a single lag when the forcing variable is endogenous. Parameter sets are $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, $\theta = \{-0.5; -0.1\}$, and $\sigma_u = \sigma_\varepsilon = 1$. The median and the MAD of the parameter distribution are computed over $N = 2000$ simulations of size $T = 100$. Estimation methods are GMM with $L = 0$ and 7 lags of $W_t = \{Y_{t-1}, Z_t\}$ and ML. The Plim of the OLS estimator is also computed with a sample of $100,000$ observations.

Table 5: This table reports Plim of GMM and ML estimators in the case of measurement error when the forcing variable is endogenous. Parameter sets are $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, $\sigma_u = \sigma_\varepsilon = 1$, and $\tau = \{0.1; 0.5; 0.9\}$. These Plims are computed with a sample of $25,000$ observations.

Table 6: This table reports Plim of GMM and ML estimators in the case of omitted dynamics when the forcing variable is endogenous. Parameter sets are $\omega_f = \{0.55; 0.75; 0.95\}$, $\rho = \{0.1; 0.5; 0.9\}$, $\beta = \{0.1; 1\}$, $\theta = \{-0.5; -0.1\}$, and $\sigma_u = \sigma_\varepsilon = 1$. These Plims are computed with a sample of $25,000$ observations.

Table 7: This table reports estimates of the hybrid Phillips curve with GMM and ML over the period 1960:I-2000:IV. Panel A corresponds to the real-ULC model, while Panel B is devoted to the output-gap model. The baseline instrument set is $W_{t-1} = \{Y_{t-1}, Z_{t-1}\}$. J-stat denotes the Hansen's statistic for the test of over-identifying restrictions, lnL is the log-likelihood and see is the standard error of estimates.

Table 8: This table reports Monte-Carlo simulations with parameters close to those of the true DGP for the hybrid Phillips curve with output gap as forcing variable. The DGP parameter is: $\omega_f = 0.4$, $\omega_b^1 = 0.5$, $\omega_b^2 = 0.1$, $\beta = 0.1$, $\rho = 0.8$, $\theta = -0.05$, and $\sigma_u^2 = \sigma_\varepsilon^2 = 1$. Panel A corresponds to the probability limits of the estimators, obtained using a sample of $10,000$ observations. Panel B corresponds to the finite-sample distribution of the estimators, obtained using 1000 samples of $T = 150$ observations. The distribution is summarized using the median and the median of absolute deviation.

Figure 1: This figure displays the data on inflation, output gap, and real ULC over the period 1960:I-2000:IV.

Figure 2: This figure displays, in the plane $\{\omega_f, \omega_b^1\}$, the area where the hybrid equation with two lags is stationary.

Table 1: Finite-sample bias of the model with a single lag

Struct. parameters		Statistic	Panel A: β=0.1								Panel B: β=1							
			GMM (L=0)		GMM (L=7)		ML		Plim OLS		GMM (L=0)		GMM (L=7)		ML		Plim OLS	
ω_f	ρ		α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b
0.55	0.1	median	0.55	0.10	0.52	0.10	0.55	0.10	0.50	0.11	0.54	1.01	0.42	1.24	0.55	1.00	0.34	1.36
		MAD	0.06	0.08	0.06	0.09	0.02	0.05			0.08	0.18	0.07	0.20	0.01	0.12		
0.55	0.5	median	0.55	0.10	0.52	0.12	0.55	0.10	0.49	0.12	0.54	1.05	0.47	1.32	0.55	1.01	0.32	1.88
		MAD	0.06	0.05	0.06	0.07	0.01	0.03			0.06	0.28	0.06	0.34	0.00	0.18		
0.55	0.9	median	0.55	0.11	0.53	0.14	0.55	0.11	0.47	0.19	0.55	1.15	0.53	1.38	0.55	1.11	0.46	2.02
		MAD	0.03	0.06	0.03	0.07	0.00	0.04			0.03	0.58	0.03	0.72	0.00	0.47		
0.75	0.1	median	0.75	0.11	0.68	0.11	0.75	0.10	0.50	0.12	0.75	1.01	0.65	1.05	0.75	1.00	0.43	1.17
		MAD	0.09	0.10	0.08	0.12	0.06	0.07			0.08	0.15	0.07	0.16	0.03	0.14		
0.75	0.5	median	0.75	0.10	0.68	0.11	0.75	0.10	0.50	0.12	0.74	1.01	0.66	1.08	0.75	1.01	0.44	1.25
		MAD	0.09	0.07	0.08	0.08	0.05	0.04			0.07	0.18	0.06	0.22	0.02	0.18		
0.75	0.9	median	0.74	0.11	0.68	0.12	0.75	0.11	0.49	0.13	0.75	1.07	0.69	1.16	0.75	1.07	0.48	1.33
		MAD	0.06	0.05	0.06	0.06	0.03	0.05			0.05	0.44	0.05	0.53	0.00	0.44		
0.95	0.1	median	0.95	0.09	0.83	0.10	0.95	0.10	0.50	0.11	0.94	1.00	0.82	1.01	0.95	1.00	0.49	1.03
		MAD	0.13	0.13	0.12	0.14	0.09	0.09			0.11	0.17	0.10	0.19	0.06	0.16		
0.95	0.5	median	0.95	0.10	0.83	0.10	0.96	0.10	0.50	0.11	0.94	1.02	0.82	1.03	0.95	1.02	0.49	1.04
		MAD	0.13	0.08	0.11	0.10	0.09	0.05			0.10	0.19	0.09	0.21	0.04	0.18		
0.95	0.9	median	0.94	0.11	0.83	0.11	0.95	0.11	0.50	0.10	0.94	1.09	0.83	1.10	0.95	1.10	0.50	1.06
		MAD	0.12	0.05	0.10	0.06	0.06	0.05			0.09	0.45	0.08	0.50	0.01	0.44		

Table 2: Analytic solution for the case of measurement error

Structural parameters			Panel A: β=0.1						Panel B: β=1					
			GMM1		GMM2		ML		GMM1		GMM2		ML	
ω_f	ρ	τ	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b
0.55	0.10	0.10	0.55	0.01	0.59	0.01	0.55	0.01	0.51	0.11	0.89	0.04	0.54	0.11
0.55	0.10	0.50	0.55	0.05	0.57	0.05	0.55	0.05	0.53	0.52	0.80	0.27	0.55	0.52
0.55	0.10	0.90	0.55	0.09	0.56	0.09	0.55	0.09	0.55	0.91	0.63	0.77	0.55	0.91
0.55	0.50	0.10	0.52	0.01	0.62	0.01	0.54	0.02	0.40	0.16	0.70	0.04	0.52	0.16
0.55	0.50	0.50	0.53	0.05	0.60	0.04	0.54	0.07	0.48	0.62	0.66	0.29	0.53	0.69
0.55	0.50	0.90	0.55	0.09	0.56	0.09	0.55	0.10	0.54	0.94	0.58	0.78	0.55	0.97
0.55	0.90	0.10	0.38	0.03	0.57	0.01	0.50	0.04	0.37	0.30	0.57	0.07	0.50	0.37
0.55	0.90	0.50	0.50	0.08	0.56	0.04	0.52	0.15	0.50	0.79	0.57	0.42	0.52	1.51
0.55	0.90	0.90	0.54	0.10	0.55	0.09	0.54	0.15	0.54	0.97	0.55	0.87	0.54	1.47
0.75	0.10	0.10	0.75	0.01	0.76	0.01	0.75	0.01	0.71	0.10	0.89	0.09	0.72	0.11
0.75	0.10	0.50	0.75	0.05	0.75	0.05	0.75	0.05	0.73	0.51	0.83	0.48	0.73	0.51
0.75	0.10	0.90	0.75	0.09	0.75	0.09	0.75	0.09	0.75	0.90	0.77	0.89	0.75	0.91
0.75	0.50	0.10	0.73	0.01	0.76	0.01	0.73	0.02	0.55	0.12	0.84	0.09	0.60	0.12
0.75	0.50	0.50	0.74	0.05	0.76	0.05	0.74	0.07	0.65	0.54	0.80	0.48	0.65	0.58
0.75	0.50	0.90	0.75	0.09	0.75	0.09	0.75	0.10	0.73	0.91	0.76	0.89	0.73	0.93
0.75	0.90	0.10	0.53	0.01	0.76	0.01	0.53	0.02	0.44	0.14	0.77	0.10	0.52	0.14
0.75	0.90	0.50	0.64	0.06	0.76	0.05	0.56	0.08	0.60	0.59	0.76	0.49	0.54	0.69
0.75	0.90	0.90	0.73	0.09	0.75	0.09	0.66	0.12	0.72	0.93	0.75	0.90	0.65	1.07
0.95	0.10	0.10	0.95	0.01	0.95	0.01	0.95	0.01	0.91	0.10	0.97	0.10	0.91	0.10
0.95	0.10	0.50	0.95	0.05	0.95	0.05	0.95	0.05	0.93	0.50	0.96	0.50	0.92	0.51
0.95	0.10	0.90	0.95	0.09	0.95	0.09	0.95	0.09	0.95	0.90	0.95	0.90	0.94	0.90
0.95	0.50	0.10	0.94	0.01	0.95	0.01	0.93	0.02	0.72	0.10	0.97	0.10	0.70	0.11
0.95	0.50	0.50	0.94	0.05	0.95	0.05	0.94	0.07	0.82	0.51	0.96	0.50	0.78	0.54
0.95	0.50	0.90	0.95	0.09	0.95	0.09	0.95	0.10	0.92	0.90	0.95	0.90	0.91	0.92
0.95	0.90	0.10	0.73	0.01	0.95	0.01	0.57	0.02	0.55	0.10	0.95	0.10	0.53	0.11
0.95	0.90	0.50	0.83	0.05	0.95	0.05	0.62	0.08	0.73	0.51	0.95	0.50	0.56	0.53
0.95	0.90	0.90	0.93	0.09	0.95	0.09	0.80	0.12	0.91	0.90	0.95	0.90	0.73	0.93
Bounds														
0.5	0.0	0.0	0.50		1.00		0.50							
0.5	1.0	0.0	0.25		0.50		0.50							
1.0	0.0	0.0	0.50		1.00		1.00							
1.0	1.0	0.0	1.00		1.00		0.50							

Table 3: Analytic solution for the case of omitted dynamics

Structural parameters					Panel A: β=0.1						Panel B: β=1					
					GMM1		GMM2		ML		GMM1		GMM2		ML	
ω_f	$\omega_b^{\ 1}$	$\omega_b^{\ 2}$	$\varphi_1+\varphi_2$	ρ	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b
0.40	1.20	-0.60	0.00	0.00	0.25	0.16	0.48	0.07	0.63	0.16	0.25	1.56	0.48	0.69	0.63	1.56
0.40	1.20	-0.60	0.00	0.50	0.20	0.31	0.48	0.18	0.59	0.15	0.04	3.81	0.49	1.79	0.51	1.07
0.40	1.20	-0.60	0.00	0.90	0.07	0.28	0.43	0.25	0.52	0.02	0.04	2.86	0.41	2.55	0.51	0.20
0.40	1.00	-0.40	0.50	0.00	0.29	0.14	0.50	0.06	0.57	0.14	0.29	1.43	0.50	0.63	0.57	1.43
0.40	1.00	-0.40	0.50	0.50	0.29	0.29	0.50	0.13	0.57	0.18	0.29	2.86	0.50	1.25	0.57	1.79
0.40	1.00	-0.40	0.50	0.90	0.29	0.37	0.50	0.23	0.57	0.10	0.29	3.66	0.50	2.34	0.57	1.01
0.40	0.84	-0.24	0.90	0.00	0.32	0.13	0.57	0.04	0.52	0.13	0.32	1.29	0.57	0.36	0.52	1.29
0.40	0.84	-0.24	0.90	0.50	0.35	0.23	0.55	0.00	0.52	0.17	0.38	1.95	0.53	0.32	0.52	1.78
0.40	0.84	-0.24	0.90	0.90	0.47	0.27	0.50	0.15	0.54	0.24	0.47	2.72	0.50	1.56	0.54	2.40
0.65	0.70	-0.35	0.00	0.00	0.48	0.11	0.54	0.11	0.74	0.11	0.48	1.14	0.54	1.10	0.74	1.14
0.65	0.70	-0.35	0.00	0.50	0.49	0.14	0.54	0.14	0.75	0.11	0.54	1.39	0.59	1.36	0.81	1.27
0.65	0.70	-0.35	0.00	0.90	0.56	0.15	0.61	0.15	0.84	0.11	0.59	1.52	0.64	1.51	0.88	1.26
0.65	0.38	-0.03	0.50	0.00	0.63	0.10	0.63	0.10	0.66	0.10	0.63	1.01	0.63	1.01	0.66	1.01
0.65	0.38	-0.03	0.50	0.50	0.64	0.10	0.64	0.10	0.66	0.10	0.65	1.03	0.65	1.03	0.66	1.03
0.65	0.38	-0.03	0.50	0.90	0.65	0.11	0.65	0.11	0.67	0.10	0.66	1.05	0.66	1.05	0.67	1.05
0.65	0.12	0.23	0.90	0.00	0.85	0.08	1.10	0.06	0.54	0.08	0.85	0.83	1.10	0.63	0.54	0.83
0.65	0.12	0.23	0.90	0.50	0.81	0.03	1.05	-0.03	0.54	0.07	0.67	0.66	0.79	0.35	0.53	0.70
0.65	0.12	0.23	0.90	0.90	0.56	0.06	0.58	0.03	0.52	0.06	0.56	0.61	0.57	0.46	0.52	0.62
0.90	0.20	-0.10	0.00	0.00	0.82	0.10	0.82	0.10	0.91	0.10	0.82	1.01	0.82	1.01	0.91	1.01
0.90	0.20	-0.10	0.00	0.50	0.82	0.11	0.82	0.11	0.91	0.10	0.85	1.06	0.85	1.06	0.94	1.05
0.90	0.20	-0.10	0.00	0.90	0.86	0.11	0.86	0.11	0.95	0.10	0.89	1.10	0.89	1.10	0.99	1.10
0.90	-0.25	0.35	0.50	0.00	1.38	0.09	1.42	0.09	0.85	0.09	1.38	0.94	1.42	0.94	0.85	0.94
0.90	-0.25	0.35	0.50	0.50	1.36	0.05	1.40	0.05	0.83	0.10	1.07	0.67	1.10	0.65	0.72	0.76
0.90	-0.25	0.35	0.50	0.90	0.92	0.04	0.94	0.03	0.67	0.07	0.81	0.56	0.82	0.54	0.64	0.61
0.90	-0.61	0.71	0.90	0.00	3.10	0.07	4.24	0.06	0.66	0.07	3.10	0.73	4.24	0.59	0.66	0.73
0.90	-0.61	0.71	0.90	0.50	2.93	-0.18	4.08	-0.31	0.64	0.07	1.44	-0.13	2.22	-1.01	0.55	0.48
0.90	-0.61	0.71	0.90	0.90	0.70	-0.04	0.87	-0.15	0.52	0.04	0.61	0.18	0.67	-0.19	0.52	0.35
Bounds																
1.00	1.0		-1.0	0.0	0.50		0.50		1.00							
1.00	1.0		-1.0	1.0	0.50		0.50		1.00							
1.00	-1.0		1.0	0.0	infinite		infinite		1.00							
1.00	-1.0		1.0	1.0	0.50		infinite		0.50							
0.33	1.0		1.0	0.0	0.25		0.50		0.50							
0.33	1.0		1.0	1.0	0.25		0.50		0.50							

Table 4: Finite-sample bias of the model with endogenous forcing variable

Struct. parameters			**Statistic**	**Panel A: β=0.1**								**Panel B: β=1**							
				GMM (L=0)		GMM (L=7)		ML		Plim OLS		GMM (L=0)		GMM (L=7)		ML		Plim OLS	
ω_f	ρ	θ		α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b
0.25	0.50	-0.50	median	0.43	0.04	0.48	0.03	0.40	0.06	0.49	0.02	0.25	1.01	0.22	1.05	0.25	1.00	0.15	1.14
			MAD	0.13	0.05	0.04	0.02	0.10	0.04			0.05	0.08	0.04	0.08	0.04	0.08		
0.25	0.75	-0.50	median	0.38	0.06	0.47	0.03	0.34	0.08	0.49	0.02	0.25	1.00	0.23	1.04	0.25	1.00	0.17	1.14
			MAD	0.11	0.04	0.04	0.02	0.09	0.04			0.04	0.07	0.03	0.07	0.03	0.07		
0.25	0.90	-0.50	median	0.33	0.08	0.43	0.04	0.30	0.09	0.48	0.02	0.25	1.01	0.23	1.03	0.25	1.00	0.18	1.13
			MAD	0.09	0.03	0.04	0.02	0.08	0.03			0.03	0.07	0.03	0.07	0.03	0.06		
0.50	0.75	-0.50	median	0.50	0.11	0.48	0.12	0.50	0.11	0.43	0.14	0.49	1.02	0.44	1.08	0.50	1.01	0.29	1.12
			MAD	0.06	0.04	0.04	0.04	0.04	0.03			0.04	0.08	0.03	0.08	0.03	0.07		
0.75	0.75	-0.50	median	0.73	0.11	0.65	0.14	0.75	0.10	0.43	0.23	0.73	1.01	0.64	1.06	0.75	1.01	0.35	1.34
			MAD	0.06	0.04	0.05	0.05	0.05	0.03			0.06	0.09	0.05	0.10	0.04	0.09		
0.25	0.50	-0.10	median	0.46	0.04	0.49	0.03	0.44	0.05	0.50	0.03	0.25	1.00	0.26	0.98	0.25	1.00	0.26	0.99
			MAD	0.11	0.05	0.03	0.03	0.06	0.04			0.06	0.13	0.05	0.12	0.05	0.12		
0.25	0.75	-0.10	median	0.47	0.03	0.49	0.02	0.42	0.05	0.49	0.02	0.25	1.01	0.25	1.02	0.25	1.00	0.24	1.03
			MAD	0.12	0.05	0.03	0.02	0.06	0.03			0.04	0.11	0.04	0.11	0.04	0.10		
0.25	0.90	-0.10	median	0.47	0.03	0.49	0.01	0.38	0.06	0.49	0.01	0.25	1.00	0.25	1.01	0.25	1.01	0.23	1.05
			MAD	0.12	0.05	0.03	0.02	0.07	0.03			0.03	0.09	0.03	0.09	0.02	0.08		
0.50	0.75	-0.10	median	0.50	0.11	0.49	0.11	0.50	0.11	0.48	0.07	0.49	1.02	0.46	1.10	0.50	1.01	0.33	1.35
			MAD	0.04	0.05	0.03	0.04	0.03	0.03			0.03	0.10	0.03	0.11	0.03	0.09		
0.75	0.75	-0.10	median	0.74	0.11	0.67	0.12	0.76	0.10	0.49	0.16	0.74	1.02	0.66	1.06	0.75	1.01	0.44	1.20
			MAD	0.06	0.04	0.05	0.04	0.03	0.02			0.05	0.12	0.04	0.13	0.03	0.11		

Table 5: Asymptotic solution for the case of measurement error in the model with endogenous forcing variable

Structural parameters				**Panel A: β=0.1**						**Panel B: β=1**					
				GMM1		GMM2		ML		GMM1		GMM2		ML	
ω_f	ρ	θ	τ	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b
0.25	0.75	-0.50	0.10	0.47	0.00	0.52	0.00	0.54	0.00	0.37	0.39	0.52	0.27	0.52	0.37
0.25	0.75	-0.50	0.50	0.43	0.02	0.50	0.01	0.47	0.03	0.28	0.85	0.33	0.77	0.31	0.83
0.25	0.75	-0.50	0.90	0.32	0.07	0.44	0.03	0.33	0.07	0.56	0.47	0.66	0.42	0.69	0.47
0.50	0.75	-0.50	0.50	0.54	0.04	0.57	0.03	0.58	0.06	0.77	0.51	0.85	0.49	0.86	0.53
0.75	0.75	-0.50	0.50	0.75	0.05	0.77	0.05	0.79	0.08	0.53	0.06	0.68	0.04	0.70	0.07
0.25	0.50	-0.50	0.10	0.51	0.00	0.54	0.00	0.54	0.00	0.44	0.37	0.58	0.27	0.55	0.36
0.25	0.50	-0.50	0.50	0.47	0.01	0.52	0.00	0.48	0.02	0.30	0.84	0.35	0.77	0.32	0.83
0.25	0.50	-0.50	0.90	0.34	0.06	0.45	0.03	0.34	0.07	0.40	0.07	0.58	0.04	0.64	0.08
0.25	0.90	-0.50	0.10	0.42	0.01	0.51	0.00	0.52	0.01	0.35	0.42	0.50	0.28	0.51	0.38
0.25	0.90	-0.50	0.50	0.37	0.03	0.49	0.01	0.43	0.04	0.27	0.87	0.32	0.78	0.30	0.84
0.25	0.90	-0.50	0.90	0.29	0.08	0.42	0.04	0.30	0.08	0.33	0.08	0.58	0.02	0.52	0.08
0.25	0.75	-0.10	0.10	0.47	0.00	0.53	0.00	0.50	0.01	0.30	0.43	0.54	0.13	0.42	0.40
0.25	0.75	-0.10	0.50	0.43	0.02	0.52	0.00	0.42	0.03	0.26	0.87	0.39	0.57	0.29	0.84
0.25	0.75	-0.10	0.90	0.35	0.06	0.46	0.03	0.33	0.07	0.46	0.55	0.60	0.40	0.47	0.67
0.50	0.75	-0.10	0.50	0.48	0.00	0.55	0.00	0.50	0.01	0.67	0.51	0.81	0.46	0.70	0.87
0.75	0.75	-0.10	0.50	0.44	0.02	0.53	0.01	0.43	0.03	0.40	0.07	0.65	0.02	0.53	0.08
0.25	0.50	-0.10	0.10	0.31	0.07	0.48	0.03	0.30	0.08	0.35	0.40	0.61	0.14	0.43	0.39
0.25	0.50	-0.10	0.50	0.43	0.00	0.51	0.00	0.49	0.01	0.28	0.85	0.43	0.56	0.29	0.84
0.25	0.50	-0.10	0.90	0.40	0.02	0.52	0.00	0.40	0.04	0.30	0.08	0.54	0.02	0.52	0.08
0.25	0.90	-0.10	0.10	0.32	0.07	0.49	0.01	0.30	0.08	0.28	0.45	0.51	0.13	0.42	0.39
0.25	0.90	-0.10	0.50	0.44	0.07	0.61	0.04	0.66	0.08	0.25	0.89	0.38	0.56	0.28	0.85
0.25	0.90	-0.10	0.90	0.49	0.01	0.51	0.00	0.42	0.04	0.44	0.22	0.51	0.13	0.42	0.39

Table 6: Asymptotic solution for the case of omitted dynamics in the model with endogenous forcing variable

Structural parameters					Panel A: β=0.1						Panel B: β=1					
					GMM1		GMM2		ML		GMM1		GMM2		ML	
ω_f	ω_b^1	ω_b^2	ρ	θ	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b	α_f	b
0.33	0.33	0.34	0.50	-0.5	1.09	-0.09	1.45	-0.18	0.57	0.00	0.75	0.70	0.73	0.72	0.55	0.77
0.33	0.33	0.34	0.75	-0.5	1.04	-0.10	1.31	-0.17	0.55	0.01	0.66	0.61	0.63	0.64	0.51	0.69
0.33	0.33	0.34	0.90	-0.5	0.86	-0.07	1.14	-0.14	0.49	0.02	0.60	0.56	0.57	0.59	0.47	0.64
0.25	0.25	0.50	0.50	-0.5	1.62	-0.18	2.03	-0.26	0.62	0.00	0.90	0.38	0.75	0.50	0.55	0.58
0.50	0.00	0.50	0.75	-0.5	1.61	-0.21	1.94	-0.30	0.68	0.01	1.24	0.59	1.00	0.72	0.77	0.78
0.50	0.25	0.25	0.90	-0.5	0.86	-0.05	0.95	-0.09	0.58	0.04	0.73	0.79	0.71	0.81	0.60	0.84
0.65	0.35	0.00	0.50	-0.5	0.65	0.10	0.65	0.10	0.65	0.10	0.66	0.99	0.66	0.99	0.65	0.99
0.75	0.75	-0.50	0.75	-0.5	0.45	0.37	0.48	0.35	0.70	0.24	0.41	1.48	0.38	1.51	0.53	1.41
0.75	-0.50	0.75	0.90	-0.5	3.99	-0.51	3.46	-0.39	0.93	-0.03	4.69	0.76	1.51	1.48	1.93	1.55
0.33	0.33	0.34	0.50	-0.1	0.94	-0.03	1.63	-0.16	0.47	0.03	0.59	0.51	0.75	0.31	0.42	0.60
0.33	0.33	0.34	0.75	-0.1	1.06	-0.10	1.58	-0.22	0.49	0.02	0.50	0.45	0.59	0.30	0.39	0.53
0.33	0.33	0.34	0.90	-0.1	0.93	-0.10	1.47	-0.24	0.47	0.02	0.43	0.46	0.48	0.37	0.37	0.52
0.50	0.00	0.50	0.75	-0.1	1.50	-0.20	2.28	-0.39	0.54	0.02	0.74	0.31	0.79	0.24	0.49	0.49
0.50	0.25	0.25	0.90	-0.1	0.80	-0.06	1.09	-0.18	0.50	0.04	0.59	0.62	0.61	0.59	0.49	0.68
0.65	0.35	0.00	0.50	-0.1	0.65	0.10	0.65	0.10	0.65	0.10	0.65	1.00	0.65	1.00	0.65	1.00
0.75	0.75	-0.50	0.75	-0.1	0.51	0.19	0.55	0.18	0.83	0.11	0.54	1.57	0.55	1.56	0.81	1.36
0.75	-0.50	0.75	0.90	-0.1	3.15	-0.48	3.98	-0.66	0.69	0.04	1.60	-0.04	1.18	0.28	0.70	0.56

Table 7: Estimates of the hybrid Phillips curve with GMM and ML (1960:I-2000:IV)

	GMM1 ($\boldsymbol{W_{t-1}}$)		**GMM2** ($\boldsymbol{W_{t-1},\ldots,W_{t-3}}$)		**ML**		**ML (3-lag model)**	
	Parameter	std dev.	Parameter	std dev.	Parameter	std dev.	Parameter	std dev.
				Panel A: Real-ULC model				
ω_f	0.645	*0.162*	0.656	*0.149*	0.504	*0.024*	0.538	*0.041*
ω_{b1}	0.345	*0.162*	0.334	*0.149*	0.486	*0.024*	0.430	*0.077*
ω_{b2}	-		-		-		-0.096	*0.089*
ω_{b3}	-		-		-		0.117	*0.049*
β	0.026	*0.020*	0.021	*0.020*	0.007	*0.003*	0.008	*0.004*
	J-stat	*p-value*	J-stat	*p-value*	lnL	see	lnL	see
	-		2.769	0.837	-426.858	1.073	-419.539	1.044
				Panel B: Output-gap model				
ω_f	0.558	*0.162*	0.677	*0.121*	0.453	*0.034*	0.396	*0.164*
ω_{b1}	0.432	*0.162*	0.313	*0.121*	0.537	*0.034*	0.471	*0.104*
ω_{b2}	-		-		-		-0.076	*0.097*
ω_{b3}	-		-		-		0.200	*0.072*
β	0.008	0.056	-0.037	0.049	0.036	0.018	0.103	*0.063*
	J-stat	*p-value*	J-stat	*p-value*	lnL	see	lnL	see
	-		5.282	0.508	-422.821	1.063	-407.494	1.006

Table 8: Monte-Carlo simulations

	GMM1 (W_{t-1})		GMM2 ($W_{t-1},\ldots,W_{t-3}$)		ML	
	Panel A: Plim of estimators					
	Plim		Plim		Plim	
ω_f	0.524		0.746		0.425	
ω_{b1}	0.466		0.244		0.565	
β	0.023		-0.077		0.060	
	Panel B: Finite-sample estimators (T=150)					
	Median	MAD	Median	MAD	Median	MAD
ω_f	0.533	0.068	0.651	0.151	0.437	0.042
ω_{b1}	0.457	0.068	0.339	0.151	0.553	0.042
β	0.024	0.040	-0.029	0.070	0.064	0.024

Figure 1: Inflation, real ULC, and output gap

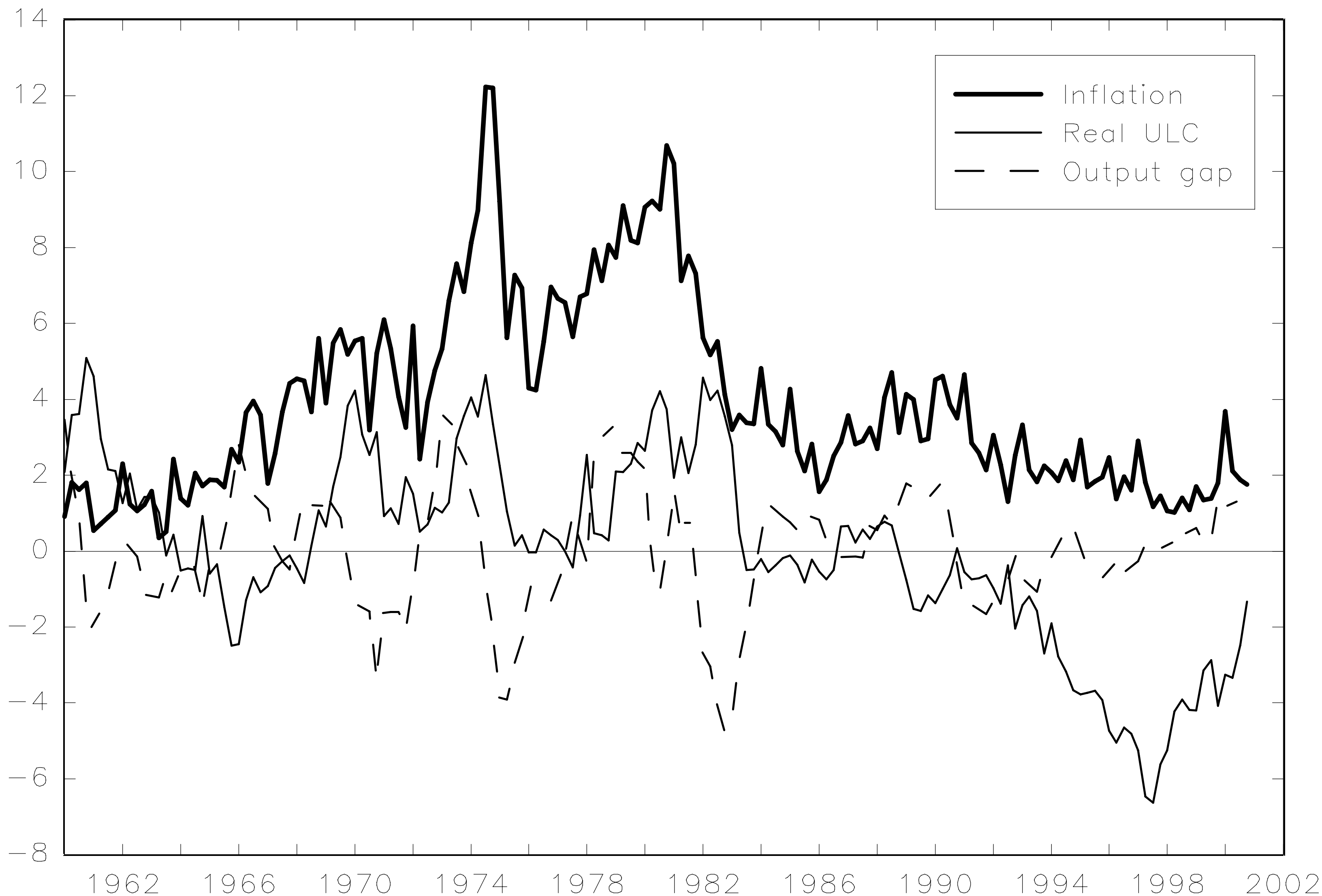
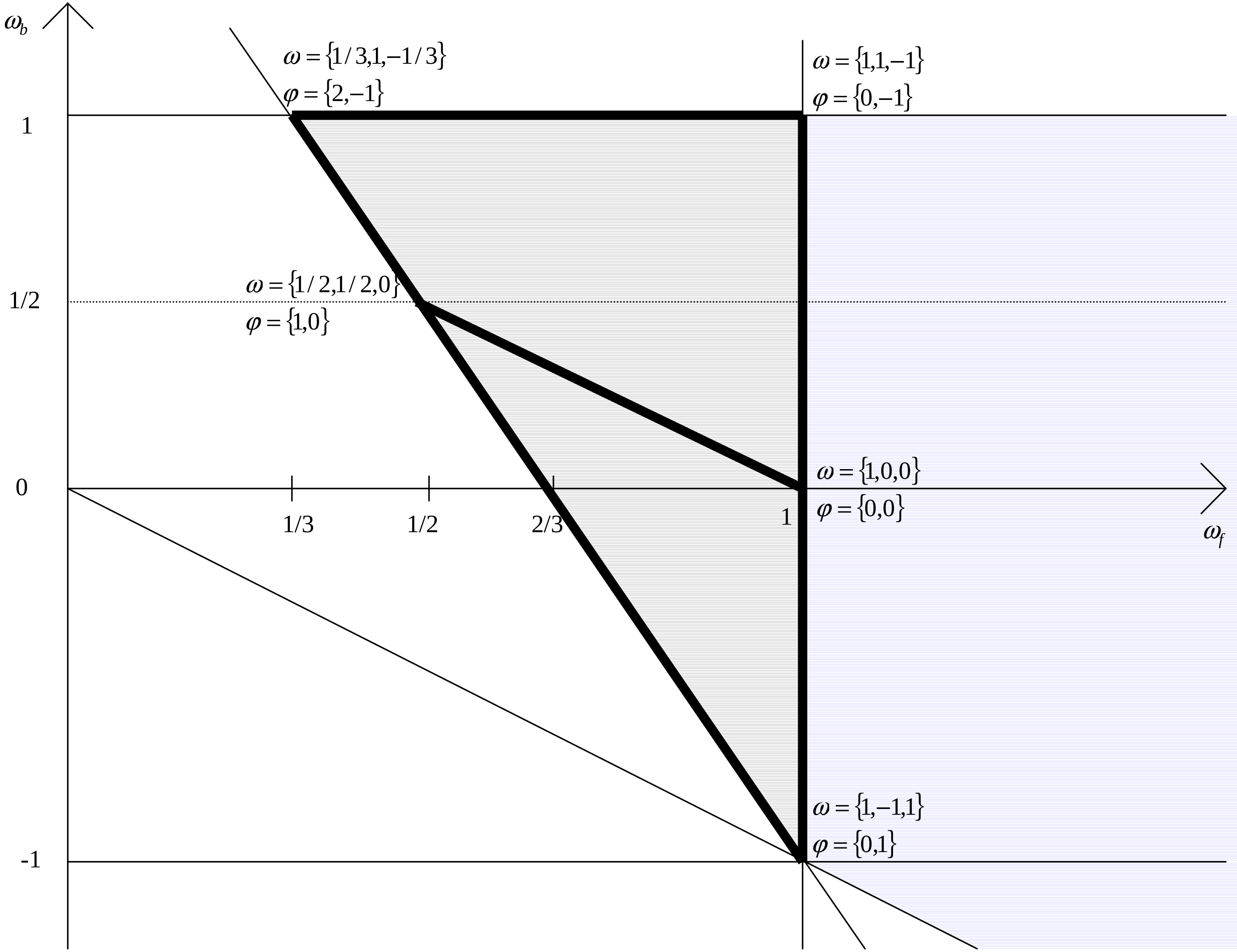


Figure 2



ω_b
$\omega = \{1/3, 1, -1/3\}$
$\varphi = \{2, -1\}$
$\omega = \{1, 1, -1\}$
$\varphi = \{0, -1\}$
1
1/2
$\omega = \{1/2, 1/2, 0\}$
$\varphi = \{1, 0\}$
0
$\omega = \{1, 0, 0\}$
$\varphi = \{0, 0\}$
1/3
1/2
2/3
1
ω_f
$\omega = \{1, -1, 1\}$
$\varphi = \{0, 1\}$
-1