ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Lindé, Jesper

Working Paper

Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach

Sveriges Riksbank Working Paper Series, No. 129

Provided in Cooperation with:
Central Bank of Sweden, Stockholm

Suggested Citation: Lindé, Jesper (2001) : Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach, Sveriges Riksbank Working Paper Series, No. 129, Sveriges Riksbank, Stockholm

This Version is available at:
https://hdl.handle.net/10419/82401

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach

Jesper Lindé*

Sveriges Riksbank Working Paper Series
No. 129
Revised version March 2005

Abstract

The New-Keynesian Phillips curve has recently become an important ingredient in monetary policy models. However, using limited information methods, the empirical support for the New-Keynesian Phillips curve appear to be mixed. This paper argues, by means of Monte Carlo simulations with a simple New-Keynesian sticky price model, that single equations methods, e.g. GMM, are likely to produce imprecise and biased estimates. Then, it is argued that estimating the model with full information maximum likelihood (FIML) is a useful way of obtaining better estimates. Finally, a version of the model used in the Monte Carlo simulations is estimated on U.S. data with FIML and although the pure forward-looking New-Keynesian Phillips curve is rejected, a version with both forward- and backward-looking components provides a reasonable approximation of U.S. inflation dynamics.

Keywords: New-Keynesian Phillips curve; Rational expectations IS-curve; Backward-looking Phillips curve; Measurement errors; Generalized method of moments; Full information maximum likelihood estimation.

JEL Classification Numbers: E52, C52, C22.

*Research Department, Sveriges Riksbank, 103 37 Stockholm, Sweden. *Telephone*: + 46 70 24 66 171. *Fax*: + 46 8 21 05 31. *E-mail*: `jesper.linde@riksbank.se`. I would like to thank Lawrence J. Christiano, Lars E.O. Svensson, Anders Vredin, an anonymous referee and and in particular the Editor Robert G. King for helpful comments and discussions. I have also benefitted from comments by seminar and conference participants at the Federal Reserve Bank of Cleveland, Sveriges Riksbank, Uppsala University, the FIEF conference "The Phillips Curve: New Theory and Evidence" (Stockholm, May 25-26, 2002), and the ECB workshop "Structural versus Reduced Form Modeling of Monetary Policy" (December 8-9, 2003). Jeremy Rudd generously provided me with the data used in this paper. Special thanks are also due to Ulf Söderström, who helped me discover an error in a previous version of the paper. The first draft of this paper was written while the author was a visiting scholar at the Research Department at the Federal Reserve Bank of Cleveland and their outstanding hospitality is gratefully acknowledged. The views expressed in the paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank.

1 Introduction

Galí and Gertler (1999) and Galí, Gertler and López-Salido (2001) estimate a "hybrid" version of the New-Keynesian Phillips curve, with both forward-looking and backward-looking elements included, with Generalized Method of Moments (GMM). They find that the pure New-Keynesian Phillips curve, where inflation is a function of expected future inflation and real marginal costs today, is a good approximation of inflation dynamics in both the U.S. and Europe. Roberts (2001), however, provides evidence against the New-Keynesian Phillips curve with only forward-looking elements using GMM, although – in contrast to Fuhrer (1997) – he finds there to be a clear role for forward-looking behavior.

As the evidence using GMM appear to be mixed, this paper first estimates a version of the New-Keynesian Phillips curve with both forward- and backward-looking elements included on US data by non-linear least squares (NLS). The estimate for the driving variable is found to have the wrong sign irrespective of whether real marginal costs or the output gap are used as the driving variable, a finding that is robust across subsamples. However, it should be emphasized that an important requirement for the NLS estimator to be unbiased and consistent is that the "pure-theory" version of the New-Keynesian Phillips curve holds.[1]

To examine the properties of single equation methods such as GMM (that have been extensively used in the literature) and NLS, I construct a simple macroeconomic model – similar to a version thoroughly studied by Clarida, Galí and Gertler (1999) – and study the properties of single equation estimation methods by means of Monte-Carlo simulations. Two cases are considered, one with no measurement errors in the data, and another where the data are assumed to be characterized by small measurement errors.[2] Interestingly, the simulations show that very

[1] The "pure-theory" version is a New-Keynesian Phillips curve without a shock included, see Kurmann (2004) for further discussion.

[2] As carefully explained in Section 3.2, there are both theoretical and statistical reasons to believe that the

small measurement errors produce biased estimates and can explain the NLS estimation results obtained on US data. The simulations also reveal two severe problems with GMM. First, even in the absence of measurement errors, the GMM estimates are severely biased in small samples towards finding inflation inertia, if there is inertia in aggregate demand and the interest rate, which seems to be the case empirically, see e.g. Fuhrer and Moore (1995a).[3] Interestingly, the size of the GMM bias is such that the estimation results may support a backward-looking Phillips curve, although the true Phillips curve is highly forward-looking, which is an opposite problem with GMM than that suggested by Rudd and Whelan (2005).[4] Second, the size of the GMM bias varies with the conduct of monetary policy, i.e. if there is a change in monetary policy, then the GMM estimates of the coefficients in the Phillips curve will also change in small samples, although the true parameters remain unchanged.[5] It should be stressed that these pitfalls with GMM are the same irrespective if using a "direct" estimation approach as in e.g. Galí and Gertler (1999) or the closed-form approach employed by Galí, Gertler and López-Salido (2005). Consequently, it is concluded that single equation estimation methods are not very informative when estimating the New-Keynesian Phillips curve on real-world data.

The paper proceeds to examine the properties of the Full Information Maximum Likelihood

data suffer from measurement errors.

[3]Note that, in the absence of measurement errors, the pitfalls discussed in this paper with GMM are small-sample problems as in Fuhrer, Moore and Schuh (1995). However, in the presence of measurement errors, the GMM bias persists even if the sample size is considerably increased as discussed in Section 3.3. Moreover, the problems identified with GMM are not due to bad normalizations since the estimated model is linear.

[4]Rudd and Whelan (2005) argue that GMM may produce support for the pure forward-looking New-Keynesian Phillips curve, although the true Phillips curve is completely backward-looking.

[5]For instance, if there is inertia in aggregate demand, an increased degree of interest rate smoothing will increase the persistence of inflation. And if GMM is used to estimate the Phillips curve, the estimated weight on forward-/backward-looking behavior will decrease/increase in small samples, although the true parameters remain unchanged.

(FIML) estimator on data with measurement errors. In contrast to the estimation results for the single equation methods, it is found that FIML does a good job in pinning down the true parameters on simulated data, confirming the findings by Fuhrer, Moore and Schuh (1995). Moreover, FIML is also shown to do a better job than the limited information methods, even if the model is severely misspecified and the measurement errors are non-normally distributed, which are common arguments against the use of FIML (see e.g. Tauchen, 1986).

The last part of the paper estimates a version of the macroeconomic model used in the simulations with FIML on exactly the same dataset as that used in the single equation estimations. Using various measures of the driving variables and the inflation rate, the share of forward-looking behavior is found to be about $0.30 - 0.50$ in both the Phillips curve and the aggregate demand relationship and highly significant, as opposed to the findings by Fuhrer (1997). A positive and significant coefficient of about 0.06 is also found for the driving variable in the Phillips curve. However, it should be acknowledged that in order to close the model with only three equations, these estimates are obtained using standard measures of the output gap, which is not typically the theory-consistent driving variable. Some recent studies that apply Maximum Likelihood techniques using theory-consistent real marginal costs as a driving variable, find a larger role for forward-looking behavior, see e.g. Adolfson et al. (2005) and Kurmann (2004), whereas Jondeau and Le Bihan (2004) obtain results similar to the ones obtained in this paper.

The structure of the paper is thus as follows. The next section presents some estimation results for the New-Keynesian Phillips curve on quarterly US data using non-linear least squares. Section 3 then presents the macroeconomic model and reports single equation estimation results on simulated data with and without measurement errors. The properties of FIML estimation on simulated data are studied in Section 4. In Section 5, the FIML estimation results on US data are presented. Some concluding remarks and tentative implications are provided in Section 6.

2 NLS estimation results on US data

Galí and Gertler (1999), Roberts (2001) and Rudd and Whelan (2005) estimate the following version of the New-Keynesian Phillips curve

$$\pi_t = \omega_f \mathrm{E}_t \pi_{t+1} + \omega_b \pi_{t-1} + \gamma y_t, \tag{1}$$

with GMM using instrumental variables to compute a proxy for $\mathrm{E}_t \pi_{t+1}$. Some authors refer to (1) with $\omega_f \approx 1$ and $\omega_b = 0$ as the New-Keynesian Phillips curve. In the following, (1) will be referred to as the New-Keynesian "hybrid" Phillips curve.

Rudd and Whelan (2005) show that GMM produces biased estimates of the parameters in (1), when instruments are used that belong to the true inflation equation. Therefore, another method is proposed to estimate the equation. By imposing rational expectations, $\pi_{t+1} = \mathrm{E}_t \pi_{t+1} + \eta_{t+1}$, we can rewrite (1) as

$$\pi_{t+1} = \frac{1}{\omega_f}\pi_t - \frac{\omega_b}{\omega_f}\pi_{t-1} - \frac{\gamma}{\omega_f} y_t + \eta_{t+1}, \tag{2}$$

where η_{t+1} is orthogonal to the information set in period t. Thus, it should be possible to obtain consistent estimates of the parameters in (1) by estimating (2) with non-linear least squares (henceforth NLS, because equation 2 is non-linear in the parameters).[6] A drawback of this estimation method, though, is that it cannot handle the case when $\omega_f = 0$, but given the previous results in the literature, this corner solution is a very unlikely outcome.

Before turning to the estimation results of (2), note that it can be written on the following form

$$\pi_t = \beta_1 \pi_{t-1} + \beta_2 \pi_{t-2} + \beta_3 y_{t-1} + \varepsilon_t, \tag{3}$$

which looks like a standard backward-looking Phillips curve. A difference between equations (2) and (3) is that the New-Keynesian hybrid Phillips curve implies that β_3 is negative whereas

[6] Note that this property of NLS crucially hinges on the absence of an error term in (1). Hamilton (1994) discusses this issue (see p. 426).

the traditional backward-looking Phillips curve, which imposes no restrictions on β_1 and β_2, suggests β_3 to be positive. So, if we estimate (2) and find ω_f to be significantly positive, a positive estimate of γ would then also be expected.

In Table 1, estimation results of equation (2) are reported for the period $1960Q1 - 1997Q4$. To investigate parameter stability, the equation pre- and post-1980 is also estimated.[7] Since Galí and Gertler (1999) have pointed out that real marginal costs should be used as the driving variable y rather than a measure of the output gap, both measures of y_t are used in Table 1. As is clearly seen, this is of no importance for the results: in both cases, γ is never significantly positive. These results are in sharp contrast to the results reported by Galí and Gertler, where the use of marginal costs instead of the output gap produced a positive and significant γ. Note also that the sum of ω_f and ω_b is not restricted to equal 1 in the estimations, but imposing such a restriction does not considerably affect the results (presumably because the estimated sum of ω_f and ω_b is very close to 1). The parameter estimates reported in Table 1 also seem to be stable over time, although the point estimates (i.e., $\hat{\omega}_f > 1$ and $\hat{\omega}_b < 0$) appear to be hard to reconcile with the underlying theory of the New-Keynesian Phillips curve.

The results in this section are not supportive for the New-Keynesian Phillips curve; rather they are more easily reconciled with the idea of a backward-looking Phillips curve. However, one aspect of the estimation results is difficult to reconcile with this interpretation: how come that the estimate of ω_f is so high, if the backward-looking Phillips curve is a better representation of the data? Possible explanations for this anomaly will be investigated in the next section.

[7] Note that the equation using non-farm business (NFB) data has also been estimated, but since the results are very similar, they are not reported in Table 1. In addition, the regression has been computed using HP-filtered output as the driving variable, but once more, the results were found to be very similar.

3 Properties of single equation estimation methods

In this section, I will set up a small macroeconomic model including the New-Keynesian Phillips curve. By estimating the New-Keynesian Phillips curve with NLS and GMM on simulated data sets from the model, the small sample properties of the single equation methods are studied by means of Monte Carlo simulations. Two cases will be considered: When all variables are measured without measurement errors and when there are small measurement errors present in the data.

3.1 The macroeconomic model

The model used consists of the following equations

$$\begin{aligned} \pi_t &= \omega_f \mathrm{E}_t \pi_{t+1} + \omega_b \pi_{t-1} + \gamma y_t, \\ y_t &= \beta_f \mathrm{E}_t y_{t+1} + \beta_b y_{t-1} - \beta_r \left(R_t - \mathrm{E}_t \pi_{t+1}\right) + \varepsilon_{y,t}, \\ R_t &= (1-\rho)\left(\gamma_\pi \pi_t + \gamma_y y_t\right) + \rho R_{t-1} + \varepsilon_{R,t}, \end{aligned} \tag{4}$$

where the shocks are allowed to follow univariate AR(1)-processes

$$\begin{aligned} \varepsilon_{y,t} &= \rho_y \varepsilon_{y,t-1} + u_{y,t}, \\ \varepsilon_{R,t} &= \rho_R \varepsilon_{R,t-1} + u_{R,t}. \end{aligned} \tag{5}$$

The first equation is the New-Keynesian hybrid Phillips curve, and the second is the aggregate demand equation. The last equation is an interest rate rule that the central bank is assumed to follow when deciding on the nominal interest rate. The model is very similar to that thoroughly studied in Clarida, Galí and Gertler (1999).[8]

[8] In the working paper version of this paper, it is shown how to solve the model given by (4) and (5) in Appendix A.

The following benchmark parameters are used in the model. $\omega_f = 0.30$, $\omega_b = 0.70$ and $\gamma = 0.13$ are adapted from estimates presented in Rudebusch (2002). For simplicity, I also set $\beta_f = 0.30$ and $\beta_b = 0.70$. The parameter β_r is taken from Rudebusch (2002) and set equal to 0.09. To get a reasonable parameterization of the monetary policy rule, $\gamma_\pi = 1.50$ and $\gamma_y = 0.50$ are chosen (see Taylor, 1993) but some interest rate smoothing is also allowed in the short-run by setting $\rho = 0.50$ (see e.g. Estrella and Fuhrer 2003, and Clarida, Galí and Gertler, 2000). To obtain values for ρ_R and σ_{ε_R}, the Fed Funds rate was used as a measure of R_t (y_t and π_t measured as in the previous section, see notes to Table 1), and the residuals $\hat{\varepsilon}_{R,t}$ were computed for the period $1987Q3 - 1999Q4$ according to the interest rate rule in (4). The resulting series for $\hat{\varepsilon}_R$ was then used to compute $\sigma_{\varepsilon_R} = 0.431$ and estimate (5) with OLS to obtain

$$\hat{\varepsilon}_{R,t} = \underset{(0.094)}{0.806}\hat{\varepsilon}_{R,t-1} + \hat{u}_{R,t}, \text{ D-W} = 1.60, \text{ B-G } \chi^2(4) = \underset{(0.58)}{2.87},$$

which suggests that the residuals ε_R in (5) are well modeled as a univariate AR(1) process with $\rho_R = 0.80$, since the p-value in the Breusch-Godfrey test for autocorrelation up to the 4th order indicates no remaining autocorrelation. ρ_y is assigned a similar value, 0.5, and σ_{ε_y} is set equal to 0.333 to make the standard error of output in the model similar to the data (about 3 percent).[9]

3.2 Estimation without measurement errors

The results when estimating (1) with GMM and (2) with NLS on simulated data are reported in Table 2.[10] In addition to the benchmark parameterization of the Phillips curve, two alternatives

[9] Note that σ_{u_y} and σ_{u_R} are set to 0.288 and 0.252, so that the unconditional standard deviations in ε_y and ε_R equal 0.333 and 0.431, respectively.

[10] In the "direct" single equation approach adopted here, GMM is equivalent to a 2SLS estimation since the Phillips curve is linear in the parameters and just identified, see Hamilton (1994) p. 418-21. The closed-form solution with GMM has also been estimated (see Galí, Gertler and López-Salido, 2005 for details), but since the results were essentially identical they are not reported below.

are considered: one where $\omega_f = 0.90$ and $\omega_b = 0.10$, and another where $\omega_f = 0.10$ and $\omega_b = 0.90$. Moreover, I also report results when estimating (1) when the actual data generating version of the Phillips curve is (1) plus a shock, $\varepsilon_{\pi,t}$. Altig, Christiano, Eichenbaum and Lindé (2002) derive such shocks to be variations in firms' mark-ups. To be moderate, we assume these shocks to be uncorrelated and have a standard deviation of 0.50, which is half the size of the estimate obtained by Rudebusch (2002).

As can be seen from Table 2, both estimation methods work very well in getting the estimated value of γ correct when there is no misspecification of the model and when ω_f is low relative to ω_b. When it comes to getting the correct estimates for ω_f and ω_b, the GMM estimation method produces a sizable bias on this sample size ($T = 200$) for high values of ω_f.[11] Moreover, in the case when the New-Keynesian Phillips curve is subject to a slight misspecification (we have added a white noise term), both NLS and GMM produce clearly upward/downward biased estimates of $\omega_f/\ \gamma$. I view this latter experiment as a variation of the misspecification problem with GMM identified by Rudd and Whelan (2005).[12]

The GMM small-sample bias in ω_f, ω_b and γ for high/low values of ω_f/ω_b arises when inflation is not intrinsically persistent, but inherits persistence in the model via inertia in the aggregate demand and the policy rule. It can be shown that GMM will only produce unbiased estimates when the intrinsic inflation persistence is about same as the persistence in the other

[11] It can be shown that the sample size must be substantially increased (to about $1,000$ observations, i.e. 250 years of data) for the small-sample bias to vanish.

[12] It should be noted that the size of the GMM biases reported in Table 2 for the misspecification case are moderated by a factor 2 if only lags are used as instruments, because R_t (or any other variable dated at t) will be correlated with the shock $\varepsilon_{\pi,t}$ and thus, is not a valid instrument. The reason for including R_t as an instrument when estimating the misspecified model in Table 2 was to provide a setting similar to that of Rudd and Whelan (2005).

variables.[13] And since Fuhrer and Moore (1995a) have shown that output and interest rates are typically best characterized as very persistent processes in the data, this result suggests that one should be very cautious when estimating New-Keynesian hybrid Phillips and demand curves using GMM, because the GMM estimation results can support a backward-looking Phillips curve, although the true Phillips curve is highly forward-looking. Ironically, this is an "inverse problem" with GMM to that identified by Rudd and Whelan (2005) (i.e. that GMM may produce estimation results that support a forward-looking specification, although the true Phillips curve is backward-looking).

Finally, it is notable that it is very difficult to parameterize the model to generate a negative γ and a high/low ω_f/ω_b when the data generating process for the inflation rate is actually the New-Keynesian hybrid Phillips curve, and when estimating the model with NLS (see Table 2). This implies that if the New-Keynesian Phillips curve in (1) is a good representation of the inflation dynamics in the US economy, the empirical findings in the previous section are hard to reconcile with the simulation results in this subsection. In the next subsection, measurement errors will be introduced in the model to see if they can account for the contradicting results in Tables 1 and 2.

3.3 Estimation with measurement errors

In the New-Keynesian Phillips curve (1), y_t can be shown to represent deviations in actual output from the "natural rate" of output in some models (see e.g. Rotemberg and Woodford, 1997). This "natural rate" is the level of output that would remain under flexible prices. However, when measuring the output gap, we typically apply an ad hoc detrending procedure – e.g. the HP filter – and hence, introduce a measurement error. Results in Orphanides (2000) suggest the measurement errors for the output gap to be considerable in real time, and if the Federal

[13] This issue is reported in detail in a working-paper version of this paper.

Reserve acts on basis of this information, the measurement errors will be transmitted into the real economy. Moreover, both the CPI- and GDP-deflator, which have typically been used in the literature, are normally associated with a considerable time-varying bias, see e.g. Bryan and Cecchetti (1993), Gordon (2000), and Gordon and Mandelkern (2001). Finally, statistical agencies often substantially revise national accounting data over time, which is an additional indication that measurement errors are likely to exist in practice.

Thus, it is here assumed that instead of observing the true output gap (y_t^*) and log price level ($\ln P_t^*$), the statistician computes

$$\begin{aligned} y_t &= y_t^* + \upsilon_{y,t}, \\ \ln P_t &= \ln P_t^* + \psi + \upsilon_{p,t}, \end{aligned}$$

where ψ is the average bias and $\upsilon_{p,t}$ is a stochastic component of the measurement bias. The measurement error in the price level implies that the inflation rate is measured as

$$\pi_t = \pi_t^* + \upsilon_{p,t} - \upsilon_{p,t-1}.$$

Although it could easily be assumed that both $\upsilon_{y,t}$ and $\upsilon_{p,t}$ are serially autocorrelated, both $\upsilon_{y,t}$ and $\upsilon_{p,t}$ will be assumed to be white noise. The reason for this is that it will turn out that even very small measurement errors considerably affect the estimation results. Since time-varying measurement errors in the output gap are likely to be more severe than for the inflation rate, $\text{var}(\upsilon_{y,t})$ is ad hoc set equal to 0.10 and $\text{var}(\upsilon_{p,t})$ equal to 0.01.

Agents in the model are assumed to behave as if they know the true output gap and the inflation rate, but the central bank is assumed to respond to y_t and π_t rather than to y_t^* and π_t^* when it sets the nominal interest rate, thus transmitting the measurement errors into the real side of the economy. Nevertheless, it is the case that the measurement errors introduced in the model have very small implications for the time series properties of the data. Standard

deviations and autocorrelations in individual series, and also cross correlations, are essentially the same.[14] Despite this, it appears from Table 3 that the results are very different when executing the same regressions as in Table 2.

For the benchmark calibration of the model, it is seen that the results when estimating (2) with NLS implies that the estimate of γ is biased downwards while ω_f is biased upwards. Interestingly, a smaller/larger value of ω_f/ω_b produces an even more upward/downward bias for ω_f/ω_b, and enables us to generate results on simulated data closely resembling the estimates on US data reported in Table 1. Therefore, if one believes that (small) measurement errors may be empirically relevant, it cannot be concluded from the estimation results in Table 1 that the New-Keynesian Phillips curve is inconsistent with the data. The only case when NLS works satisfactorily is when ω_f is near unity; but as ω_f becomes lower, NLS starts producing severely biased (and inconsistent) estimates. The GMM results are similar to those in Table 2, with the exception that the biases have become larger for the case of high/low ω_f/ω_b. Moreover, the reduction in the GMM bias is much smaller as the sample size is increased, as compared to when there are no measurement errors.[15]

The question then arises why these small measurement errors have these large effects on

[14] For instance, in the benchmark calibration, the standard deviations for y, π and R change from {2.946, 4.180, 6.496} to {2.964, 4.183, 6.500} percent and the autocorrelation coefficients from {.967, .980, .979} to {.960, .980, .979}. Moreover, the contemporaneous correlations of y and R with π change from {.239, .996} to {.238, .996}. Naturally, the importance of the measurement errors is dependent on the model parameterization, but for the different parameterizations considered in Table 3, the measurement errors always have a small impact on the time series properties of individual series.

[15] For example, if the sample size is increased to $1,000$, the GMM estimates for the case when $\omega_f = .90/\omega_b = .10$ are about .60 (ω_f), .40 (ω_b), and .022 (γ), so the GMM bias is not much moderated. Increasing the sample size to as much as $10,000$ observations, the mean GMM estimates are about .77 (ω_f), .23 (ω_b), and .085 (γ), so that it takes a considerable number of observations to shrink the GMM bias in case of measurement errors.

the estimation results for NLS as ω_f becomes smaller. The reason is that when estimating (2) when ω_f is close to unity, the estimated relationship has approximately a random walk component plus the forcing variable and a shock term, which implies that the small measurement errors have no greater impact. But as ω_f becomes smaller, the second difference of inflation is essentially being related to the forcing variable plus a shock term, and the results will be very sensitive to measurement errors in both the driving variable and the dependent variable, since the "dependent variable" is a very noisy process even without measurement errors.

To conclude this section, it seems very difficult to obtain reliable estimates of the parameter in the New-Keynesian hybrid Phillips curve using single equation estimation methods. Neither NLS estimation of (2) nor direct GMM estimation of (1) work well on sample sizes that have typically been used in empirical studies, if very small measurement errors exist in the data. Even in a world without measurement errors, the estimation results in Table 2 indicated severe problems with GMM.

4 Better estimates via FIML

In this section, the Full Information Maximum Likelihood (FIML) method will be applied to examine if this is a possible way of obtaining good estimates of the parameters in (4).

To investigate the properties of FIML through simulations, the number of shocks and variables in the model must be the same, and a shock is therefore added to the New-Keynesian hybrid Phillips Curve in the model (4)

$$\pi_t = \omega_f \mathrm{E}_t \pi_{t+1} + \omega_b \pi_{t-1} + \gamma y_t + \varepsilon_{\pi,t}, \tag{6}$$

and allow $\varepsilon_{\pi,t}$ to follow a univariate AR(1)-process,

$$\varepsilon_{\pi,t} = \rho_\pi \varepsilon_{\pi,t-1} + u_{\pi,t}$$

where we set $\sigma_{\varepsilon_\pi} = 1.012$ (taken from Rudebusch, 2002) and $\rho_\pi = 0$. Nevertheless, ρ_π will be included as a parameter in the FIML estimations.[16]

The FIML estimation results are presented in Table 5. As in Sections 3.2 and 3.3, results for different values of ω_f and ω_b are presented to see if FIML can accurately distinguish between a completely forward- and backward specification of the Phillips curve. Furthermore, results are also presented when the measurement errors are non-normally distributed, and the model is misspecified.

Interestingly, it is seen from Table 4 that FIML can discriminate between forward- and backward-looking specifications of the Phillips curve, although there are measurement errors in the data. Only for the case when ω_f/ω_b is low are the estimates somewhat biased on this sample size.[17] This is in sharp contrast to the estimation results with the single equation methods in the previous section.

Although not reported in Table 4, the FIML estimations have also been executed using the completely backward-looking Phillips curve in (3) instead of (6) as part of the data generating process, to see if the conclusion that FIML can distinguish between forward- and backward-looking specifications of the Phillips curve holds in a more general setting. After estimating (3) on quarterly US data, $\beta_1 = 0.75$, $\beta_2 = 0.25$ and $\beta_3 = 0.13$ were used when generating data. In this case, the FIML estimation results (with measurement errors) of the hybrid New-Keynesian Phillips are $\omega_f = \underset{(0.102)}{0.039}$, $\omega_b = \underset{(0.102)}{0.961}$ and $\gamma = \underset{(0.029)}{0.106}$. Once more, the FIML procedure will tend to accurately accept the backward-looking specification and reject forward-looking behavior in

[16] In Appendix B, which is available in a working paper version of this paper, it is outlined how the FIML estimations have been carried out.

[17] This finding is very interesting because once a shock in the New-Keynesian hybrid Phillips curve (the error term ε_π in equation 6) is allowed for, which is the case here, neither NLS nor GMM can distinguish between the two specifications (see the "misspecification" results in Section 3.2).

the Phillips curve, and the estimated parameters in the other equations were not heavily biased from the true parameters, with the exception of β_r which was, on average, found to be lower (0.017) than the true value (0.09).

A potential major disadvantage of FIML is the assumption of normally distributed residuals, whereas a single equation method such as GMM does not require this assumption. In Table 4, the quantitative impact on the estimation results of non-normally distributed measurement errors is explored.[18] As can be seen from the Table, the estimation results for both the Phillips curve and the aggregate demand equations are still close to their true values, although the effective variances in the measurement errors are much higher in this case. The estimation results for the policy rule are somewhat inaccurate, however, because of the assumption that monetary policy responds to the observed inflation and output gap rather than to their true values.

Finally, it should be emphasized that FIML works well even if the rest of the model is severely misspecified. This was examined in the following way. Data were generated using the benchmark parameters in the model, and then the model (incorrectly) was estimated, assuming that the shocks are white noise and that there is no smoothing in the monetary policy rule (i.e. ρ, ρ_y, and ρ_R were constrained to be 0, although they equal .50, .50, and .80, respectively, in the data generating process). The estimation results of this experiment are reported last in Table 4. In comparison with the benchmark estimates, the estimation results for aggregate demand and the monetary policy rule are clearly distorted, because those equations are heavily misspecified.

[18] Specifically, I use a uniform distribution random number generator (fat tails) where the lower and upper bounds are approximated with the lowest and highest value generated by simulating a normal distribution $2,000,000$ times. Then, the resulting distribution is scaled with the assumed variance for the measurement errors. Through this procedure, the effective variance for the measurement errors in π_t and y_t becomes 0.91 and 2.89, respectively, which is substantially higher than in the benchmark case (0.01 and 0.10, respectively, see Section 3.3).

In comparison with the results in Table 3, the estimation results for the Phillips curve are still clearly superior to the NLS estimates. And since Tables 2 and 3 revealed that GMM has serious drawbacks in other respects, the results in Table 4 also suggest that FIML outperforms GMM. Therefore, it can be concluded that applying FIML on real world data appears to be a good strategy in empirical work.

5 Estimation on US data with FIML

In this section, the following model is estimated

$$\begin{aligned} \pi_t &= \omega_f \mathrm{E}_t \pi_{t+1} + (1-\omega_f)\,\pi_{t-1} + \gamma y_t + \varepsilon_{\pi,t}, \\ y_t &= \beta_f \mathrm{E}_t y_{t+1} + \left(1-\beta_f\right) \Sigma_{i=1}^{4} \beta_{y,i} y_{t-i} - \beta_r \left(R_t - \mathrm{E}_t \pi_{t+1}\right) + \varepsilon_{y,t}, \\ R_t &= \left(1 - \Sigma_{i=1}^{3} \rho_i\right) \left(\gamma_\pi \pi_t + \gamma_y y_t\right) + \Sigma_{i=1}^{3} \rho_i R_{t-i} + \varepsilon_{R,t} \end{aligned} \tag{7}$$

with FIML on quarterly US data $1960Q1 - 1997Q4$.[19] The additional lags in the aggregate demand equation and the monetary policy rule are required to make ε_y and ε_R white noise. In Table 5, the estimation results are reported using both GDP and NFB data.[20]

[19] As driving variable in the FIML estimations, the output-gap is used as a proxy for real marginal costs in the Phillips curve, since I want to close the model with only three equations (interest rate rule, aggregate demand and the pricing equation). Looking at the single equation estimation results in Table 1, my choice of using the output gap as a proxy for real marginal costs in the FIML estimations does not seem critical, given the similarity of the estimation results when either the output gap or the real marginal costs are used as driving variables. All variables are demeaned prior to estimation.

[20] As starting values in the estimations, the benchmark parameters adapted in Sections 3 and 4 for the Phillips curve and the aggregate demand equation are used, whereas the starting values for the policy rule parameters are obtained by a simple OLS regression for the relevant sample period. It should be noted that depending on the starting values, FIML can converge to local equilibria with more or less plausible parameters. Finally, as in the previous subsection, $\sigma^2_{u_y}$, $\sigma^2_{u_\pi}$ and $\sigma^2_{u_R}$ are not included in the FIML maximization procedure, but the implied estimates (see Appendix B) are reported in Table 5.

From Table 5, it can be seen that the estimation results are similar regardless of whether GDP or NFB data are used, except for the estimates of ω_f. The mean log-likelihood function is somewhat higher when GDP data are used. The estimated standard deviations $\sigma^2_{u_\pi}$, $\sigma^2_{u_y}$ and $\sigma^2_{u_R}$, suggest that the model's fit for π and y is somewhat better using GDP rather than NFB data.

Regarding the parameter estimates, it is now seen that backward-looking behavior seems more important than forward-looking behavior, and that γ is now positive and highly significant, in contrast to the estimation results in Table 1. But if acknowledging the existence of (small) measurement errors, the differences in the estimates of ω_f reported in Tables 1 and 5 were to be expected, according to the analysis in Section 3.3 (see Table 3). So unlike the findings by Galí and Gertler (1999), the FIML estimation results suggest that the pure forward-looking New-Keynesian Phillips curve is not an acceptable approximation of US inflation dynamics. This difference is completely consistent with the findings by Rudd and Whelan (2005) and the simple misspecification problem discussed in Section 3.2. In accordance with Fuhrer and Moore (1995b), inflation inertia seems important, a result which is independent of whether GDP or NFB data are considered. Interestingly, the estimate of ω_f on GDP data is very similar to that obtained by Rudebusch (2002) using a survey measure of $\mathrm{E}_t\pi_{t+1}$.

The estimate for ω_f on GDP data is also similar to the benchmark estimate obtained by Fuhrer (1997) (0.20), but unlike Fuhrer's, the estimate here is highly significant, implying that forward-looking behavior is important for price setting behavior.[21] The estimated degree of

[21] It should, however, be noted that the log-likelihood is very flat around the optimum. Using a Likelihood Ratio (LR-) test as in Fuhrer (1997), it is then not possible to reject the joint null hypotheses that $\omega_b = 1$ and $\omega_f = 0$. But using other tests, e.g. the Wald-test, the null hypothesis that $\omega_b = 1$ and $\omega_f = 0$ is firmly rejected. In fact, the likelihood function is so flat around the optimum that it turns out that according to the LR-test, the null hypothesis that all variables are univariate AR(1)-processes cannot be rejected. But since the model (see

forward-looking in the aggregate demand equation is roughly of the same size as in the Phillips curve. The estimates of the real interest rate elasticity in the aggregate demand equation are very similar to that reported in Rudebusch (2002).

Finally, it should be emphasized that the main features of the estimation results are not sensitive to the used measure of the output gap. If HP-filtered GDP/NFB data are instead used (smoothness coefficient λ set to 1600), the main differences are that $\hat{\gamma}$ increases to about 0.12 for both GDP and NFB data, and that $\hat{\omega}_f$ for NFB data falls to about 0.30.

6 Concluding remarks

Based on Monte-Carlo simulations with a standard New-Keynesian sticky price model, it is argued in this paper that it is very difficult to obtain reliable estimates of the parameters in the New-Keynesian hybrid Phillips curve using single equation methods. In contrast to limited information methods, the FIML estimator can accurately distinguish between a forward- or backward-looking specification of the Phillips curve. Even in the presence of severe measurement errors and model misspecification, FIML still seems superior to single equation methods.

To demonstrate the usefulness of the FIML approach, the simple New-Keynesian sticky price model is estimated on quarterly US postwar data with both FIML and a limited information method. In contrast to estimation results using limited information methods, the FIML results suggest a "hybrid" version of the New-Keynesian Phillips curve where backward-looking behavior

footnote 13) implies that each variable is highly autocorrelated for a parameterization similar to the estimation results in Table 5, and thus well statistically described as univariate AR(1) processes, it is the case that the power of the LR-test used by Fuhrer is most likely very low, whereas other tests (such as the Wald-test) ought to have higher power. Therefore, I draw the conclusion that the Wald-test ought to be more informative that the LR-test used by Fuhrer.

is about equally or more important than forward-looking behavior in order to fit the data well.[22] Given the problems with single equation methods identified in this paper, this difference is not surprising. At the same time, it should be acknowledged that the estimates are obtained using a standard measure of the output gap, which is typically not the theory-consistent driving variable.

In contrast to the results reported by Fuhrer (1997), I find forward-looking behavior to be significant, but this conclusion rests upon which test is used. If a Likelihood Ratio (LR-) test is used (as Fuhrer did), then the null hypothesis of no forward-looking behavior in the Phillips curve cannot be rejected. Using another test – the Wald-test – the null of no forward-looking behavior is, in contrast, rejected. Since the New-Keynesian Phillips curve including a substantial degree of forward-looking behavior implies a very persistent process for the inflation rate as long as the driving variable is persistent, the LR-test will presumably have low power to distinguish between the null and the alternative in comparison to the Wald-test. Empirically, it seems to be the case that the labor share and the output gap are highly persistent processes. Therefore, it is argued that we should trust the results of the Wald-test which suggest a clear role for both forward- and backward-looking behavior in U.S. inflation dynamics.

[22] The estimation results are well in line with the version of the New-Keynesian with dynamic indexation shown by Altig et al. (2002) to fit the impulse response functions to monetary policy and permanent technology shocks well.

References

Adolfson, M., S. Laséen, J. Lindé, and M. Villani, 2005, The Role of Sticky Prices in An Open Economy DSGE Model: A Bayesian Investigation, *Journal of the European Economic Association Papers and Proceedings*, forthcoming.

Altig, D.E., L.J. Christiano, M. Eichenbaum, and J. Lindé, 2002, Technology shocks and aggregate fluctuations, manuscript, Northwestern University.

Bryan, M.F. and S.G. Cecchetti, 1993, The Consumer Price Index as a measure of inflation", NBER working paper 4505.

Clarida, R., J. Galí and M. Gertler, 1999, The science of monetary policy: A new Keynesian perspective, *Journal of Economic Literature* 37(4), 1661-1707.

Clarida, R., J. Galí and M. Gertler, 2000, Monetary policy rules and macroeconomic stability: Evidence and some theory, *Quarterly Journal of Economics* 115(1), 147-180.

Estrella, A. and J.C. Fuhrer, 2003, Monetary policy shifts and the stability of monetary policy models, *Review of Economics and Statistics* 85(1), 94-104.

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

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.

Fuhrer, J.C. and G.R. Moore, 1995a, Monetary policy trade-offs and the correlation between nominal interest rates and real output, *American Economic Review* 85(1), 219-239.

Fuhrer, J.C. and G.R. Moore, 1995b, Inflation persistence, *Quarterly Journal of Economics* 110(1), 127-159.

Galí, J. and M. Gertler, 1999, Inflation dynamics: A structural econometric analysis, *Journal of Monetary Economics* 44(2), 195-222.

Galí, J., M. Gertler, and J.D. López-Salido, 2001, European inflation dynamics, *European Economic Review* 45(7), 1237-1270.

Galí, J., M. Gertler, and J.D. López-Salido, 2005, Robustness of the estimates of the hybrid new Keynesian Phillips curve, *Journal of Monetary Economics*, this issue.

Gordon, R.J., 2000, Apparel prices and the Hulten/Bruegel paradox, manuscript, Northwestern University.

Gordon, R.J. and G. Mandelkern, 2001, Local indexes of apartment rent and house sale prices, manuscript, Northwestern University.

Hamilton, J.D., 1994, *Time Series Analysis* (Princeton University Press, Princeton, New Jersey).

Jondeau. E. and H. Le Bihan, 2004, Testing for the New Keynesian Phillips Curve: Additional International Evidence, *Economic Modelling*, forthcoming.

Kurmann, A., 2004, Maximum Likelihood Estimation of Dynamic Stochastic Theories with an Application to New Keynesian Pricing, manuscript, Universié du Québec à Montréal and CIRPÉE.

Orphanides, A., 2000, Activist stabilization policy and inflation: The Taylor rule in the 1970s, Finance and Economics Discussion Series 2000-13, Federal Reserve Board.

Roberts, J.M., 2001, How well does the new Keynesian sticky-price model fit the data?, Finance and Economics Discussion Series 2001-13, Federal Reserve Board.

Rotemberg, J.J. and M. Woodford, 1997, An optimization-based econometric framework for the evaluation of monetary policy, in: B.S. Bernanke and J.J. Rotemberg, eds., *NBER Macroeconomics Annual* (MIT Press, Cambridge).

Rudd, J. and K. Whelan, 2005, New tests of the new-Keynesian Phillips curve, *Journal of Monetary Economics*, this issue.

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

Tauchen, G., 1986, Statistical properties of Generalized Method-of-Moments estimators of structural parameters obtained from financial market data, *Journal of Business and Economic Statistics* 4, 397-425.

Taylor, J.B., 1993, Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.

Table 1: NLS estimation results of the New-Keynesian Phillips curve (2).

	y_t = output gap					y_t = real marginal costs				
	$\hat{\omega}_f$	$\hat{\omega}_b$	$\hat{\gamma}$	R^2	D-W	$\hat{\omega}_f$	$\hat{\omega}_b$	$\hat{\gamma}$	R^2	D-W
Whole sample	1.433	−0.280	−0.036	0.83	2.11	1.341	−0.212	−0.015	0.82	2.05
	(0.164)	(0.142)	(0.012)			(0.147)	(0.129)	(0.017)		
pre−1980	1.397	−0.220	−0.042	0.82	2.07	1.285	−0.175	0.004	0.80	2.04
	(0.216)	(0.183)	(0.018)			(0.188)	(0.169)	(0.027)		
post−1980	1.543	−0.401	−0.024	0.84	2.20	1.544	−0.297	−0.036	0.84	2.12
	(0.278)	(0.245)	(0.019)			(0.278)	(0.226)	(0.028)		

Notes: Standard errors in parentheses. Whole sample estimation period is $1960Q1 - 1997Q4$, which implies that $T = 152$ observations are used. D-W denotes the Durbin-Watson statistic. Data have been demeaned prior to estimation. π_t is measured as $\ln(P_t/P_{t-1})$ where P_t is the GDP-deflator. The output gap is defined as the quadratically detrended log of GDP, $\ln(GDP_t) - c_0 - c_1 Trend_t - c_2 Trend_t^2$, and real marginal costs are measured as the percentage deviations from the sample mean of the labor income share of output $(\bar{S})$, that is $(S_t - \bar{S})/\bar{S}$, as in Galí and Gertler (1999).

Table 2: Estimation results of the New-Keynesian Phillips curve in (4).[a]

Model parameterization	Estimation method					
	NLS			GMM		
	$\hat{\omega}_f$	$\hat{\omega}_b$	$\hat{\gamma}$	$\hat{\omega}_f$	$\hat{\omega}_b$	$\hat{\gamma}$
Benchmark	0.301 (0.008)	0.699 (0.008)	0.130 (0.005)	0.300 (0.008)	0.700 (0.008)	0.131 (0.005)
$\omega_f = 0.90, \omega_b = 0.10$	0.902 (0.055)	0.098 (0.055)	0.144 (0.034)	0.638 (0.377)	0.362 (0.377)	0.056 (0.134)
$\omega_f = 0.10, \omega_b = 0.90$	0.100 (0.007)	0.900 (0.007)	0.130 (0.002)	0.098 (0.007)	0.902 (0.007)	0.131 (0.002)
Misspecification[b]	0.889 (0.084)	0.111 (0.084)	−0.223 (0.051)	0.594 (0.099)	0.406 (0.099)	−0.040 (0.056)

Notes: The estimates are based on simulated data generated with the model given by equations (4) and (5) , where $\omega_f = 0.30$, $\omega_b = 0.70$ and $\gamma = 0.13$ in the benchmark and misspecification cases. The numbers in parentheses are the standard errors of the simulated distribution of $10,000$ estimates. $T = 200$ observations are used in each estimation. Each simulated sample is initiated with 100 additional observations to get a stochastic initial state, and are then discarded. The NLS estimation results are based on estimating (2) and the GMM/2SLS results on estimating (1) using π_{t-1}, y_{t-1} and R_{t-1} as instruments. [a] In the estimations $\hat{\omega}_f + \hat{\omega}_b = 1$ is restricted. [b] In this case, an error term $\varepsilon_{\pi,t}$ has been added in the Phillips curve, which is assumed to be i.i.d. $N(0, \sigma_{\varepsilon_\pi})$ where σ_{ε_π} equals 0.50, and the GMM instruments are R_t, π_{t-1}, y_{t-1} and R_{t-1}.

Table 3: Estimation results of the New-Keynesian Phillips curve in (4) on simulated data with small measurement errors.[a]

Model parameterization	Estimation method					
	NLS			GMM		
	$\hat{\omega}_f$	$\hat{\omega}_b$	$\hat{\gamma}$	$\hat{\omega}_f$	$\hat{\omega}_b$	$\hat{\gamma}$
Benchmark	0.691 (0.144)	0.309 (0.144)	−0.103 (0.083)	0.300 (0.010)	0.700 (0.010)	0.131 (0.006)
$\omega_f = 0.90, \omega_b = 0.10$	0.972 (0.067)	0.028 (0.067)	0.148 (0.035)	0.562 (0.215)	0.438 (0.215)	0.016 (0.086)
$\omega_f = 0.10, \omega_b = 0.90$	2.397 (0.313)	−1.397 (0.313)	−0.594 (0.099)	0.123 (0.068)	0.877 (0.068)	0.123 (0.021)

Notes: The estimates are based on simulated data generated with the model given by (4) and (5) where $\omega_f = 0.30$, $\omega_b = 0.70$ and $\gamma = 0.13$ in the benchmark case. The numbers in parentheses are the standard errors of the simulated distribution of $10,000$ estimates. $T = 200$ observations are used in each estimation. Each simulated sample is initiated with 100 additional observations to get a stochastic initial state, and these are then discarded. The NLS estimation results are based on estimating (2) and the GMM/2SLS results on estimating (1) directly using π_{t-3}, y_{t-2}, y_{t-3}, and R_{t-3} as instruments. [a] In the estimations, $\hat{\omega}_f + \hat{\omega}_b = 1$ is restricted.

Table 4: Full information Maximum Likelihood estimation results on simulated data with measurement errors.

ω_f	ω_b	γ	β_f	β_b	β_r	γ_π	γ_y	ρ	ρ_y	ρ_π	ρ_R
FIML estimates of parameter											
0.305	0.695	0.128	0.252	0.748	0.109	1.508	0.454	0.449	0.483	−0.01	0.768
(0.062)	(0.062)	(0.038)	(0.086)	(0.086)	(0.036)	(0.044)	(0.246)	(0.138)	(0.068)	(0.072)	(0.060)
True values in model											
0.30	0.70	0.13	0.30	0.70	0.09	1.50	0.50	0.50	0.50	0.00	0.80
Sensitivity analysis: $\omega_f = 0.90$, $\omega_b = 0.10$ in model, FIML estimates											
0.892	0.108	0.148	0.190	0.810	0.133	1.449	0.394	0.482	0.505	−0.037	0.777
(0.103)	(0.103)	(0.077)	(0.179)	(0.179)	(0.092)	(0.323)	(0.506)	(0.059)	(0.097)	(0.129)	(0.053)
Sensitivity analysis: $\omega_f = 0.10$, $\omega_b = 0.90$ in model, FIML estimates											
0.208	0.792	0.099	0.157	0.843	0.148	1.545	0.449	0.386	0.486	−0.003	0.759
(0.173)	(0.173)	(0.057)	(0.151)	(0.151)	(0.066)	(0.129)	(0.153)	(0.181)	(0.071)	(0.072)	(0.067)
Non-normally distributed measurement errors, FIML estimates											
0.351	0.649	0.100	0.340	0.660	0.074	1.603	0.195	0.100	0.483	−0.009	0.768
(0.100)	(0.100)	(0.056)	(0.219)	(0.219)	(0.172)	(0.115)	(0.209)	(0.179)	(0.068)	(0.072)	(0.060)
Model misspecification: ρ, ρ_y, and ρ_R constrained to 0, FIML estimates											
0.417	0.583	0.102	0.543	0.457	0.018	1.185	−0.312	0.000	0.000	0.000	0.000
(0.053)	(0.053)	(0.022)	(0.049)	(0.049)	(0.007)	(0.083)	(0.144)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: The estimates are based on simulated data generated with the aggregate demand equation and interest rate rule in (4) and (5), and the New-Keynesian Phillips given by equation (6). The numbers in parentheses are the standard errors of the simulated distribution of $10,000$ estimates. $T = 200$ observations is used in each estimation. Each simulated sample is initiated with 100 additional observations to get a stochastic initial state, and these are then discarded. In the estimations, I restrict $\hat{\omega}_f + \hat{\omega}_b = 1$ and $\hat{\beta}_f + \hat{\beta}_b = 1$. True parameters are used as starting values in the estimations.

Table 5: Full Information Maximum Likelihood estimation results for the model in (7) on US data 1960Q1-1997Q4.

	Parameter											
Data	ω_f	γ	β_f	β_r	$\beta_{y,1}$	$\beta_{y,2}$	$\beta_{y,3}$	γ_π	γ_y	ρ_1	ρ_2	ρ_3
GDP	0.282	0.048	0.430	0.087	1.275	−0.253	0.012	0.895	0.995	1.046	−0.457	0.330
	(0.057)	(0.005)	(0.017)	(0.083)	(0.069)	(0.072)	(0.056)	(0.126)	(0.526)	(0.078)	(0.123)	(0.087)
NFB	0.457	0.048	0.425	0.156	1.310	−0.229	−0.011	0.893	1.037	1.079	−0.445	0.307
	(0.065)	(0.007)	(0.027)	(0.016)	(0.174)	(0.279)	(0.037)	(0.364)	(0.170)	(0.210)	(0.124)	(0.159)

Notes: Standard errors in parentheses. Inflation is measured as $4\ln\left(\frac{P_t}{P_{t-1}}\right)$. The sample period implies that $T = 152$ observations are used in the estimations. In the estimations, $\Sigma_{i=1}^{4}\beta_{y,i}$ in the aggregate demand equation is restricted to unity. The mean log-likelihood function, $\ln \ell$, is 9.9968 and 9.3903 for GDP and NFB data. The computed standard deviations (in percent) for the innovations $\sigma^2_{u_\pi}$, $\sigma^2_{u_y}$ and $\sigma^2_{u_R}$ are {0.79570.4006 0.7724} and {1.0591 0.6925 0.7652} for GDP and NFB data, respectively.

Appendix A Solving the model

Consider the model

$$\begin{aligned}
\pi_t &= \omega_f \mathrm{E}_t \pi_{t+1} + \omega_b \pi_{t-1} + \gamma y_t + \varepsilon_{\pi,t}, \\
y_t &= \beta_f \mathrm{E}_t y_{t+1} + \beta_b y_{t-1} - \beta_r \left(R_t - \mathrm{E}_t \pi_{t+1}\right) + \varepsilon_{y,t}, \\
R_t &= (1-\rho)\left(\gamma_\pi \pi_t + \gamma_y y_t\right) + \rho R_{t-1} + \varepsilon_{R,t}
\end{aligned} \tag{A.1}$$

where

$$\begin{aligned}
\varepsilon_{\pi,t} &= \rho_\pi \varepsilon_{\pi,t-1} + u_{\pi,t}, \\
\varepsilon_{y,t} &= \rho_y \varepsilon_{y,t-1} + u_{y,t}, \\
\varepsilon_{R,t} &= \rho_R \varepsilon_{R,t-1} + u_{R,t}.
\end{aligned} \tag{A.2}$$

We want to rewrite the model given by (A.1) and (A.2) on the form

$$A_0 \begin{bmatrix} x_{1,t+1} \\ \mathrm{E}_t x_{2,t+1} \end{bmatrix} = A_1 \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + B_1 R_t + \begin{bmatrix} u_{t+1} \\ 0 \end{bmatrix} \tag{A.3}$$

where $x_{1,t}$ is vector of predetermined variables, $x_{2,t}$ is a vector of forward-looking variables, R_t is the control variable, and u_t is a vector with shocks. Note that the AS and AD curves can be rewritten as

$$\begin{aligned}
\omega_f \mathrm{E}_t \pi_{t+1} &= \pi_t - \omega_b \pi_{t-1} - \gamma y_t - \varepsilon_{\pi,t}, \\
\beta_r \mathrm{E}_t \pi_{t+1} + \beta_f \mathrm{E}_t y_{t+1} &= y_t - \beta_b y_{t-1} + \beta_r R_t - \varepsilon_{y,t}.
\end{aligned}$$

By defining $x_{1t+1} = [\ \varepsilon_{\pi,t+1} \quad \varepsilon_{y,t+1} \quad \varepsilon_{R,t+1} \quad \pi_t \quad y_t \quad R_t\ ]'$, $\mathrm{E}_t x_{2t+1} = [\ \mathrm{E}_t \pi_{t+1} \quad \mathrm{E}_t y_{t+1}\ ]'$, and $u_{t+1} = [\ u_{\pi,t+1} \quad u_{y,t+1} \quad u_{R,t+1}\ ]'$, we have

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \omega_f & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \beta_r & \beta_f \end{bmatrix}}_{\equiv A_0} \begin{bmatrix} \varepsilon_{\pi,t+1} \\ \varepsilon_{y,t+1} \\ \varepsilon_{R,t+1} \\ \pi_t \\ y_t \\ R_t \\ \mathrm{E}_t \pi_{t+1} \\ \mathrm{E}_t y_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \rho_\pi & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho_y & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_R & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & -\omega_b & 0 & 0 & 1 & -\gamma \\ 0 & -1 & 0 & 0 & -\beta_b & 0 & 0 & 1 \end{bmatrix}}_{\equiv A_1} \begin{bmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{y,t} \\ \varepsilon_{R,t} \\ \pi_{t-1} \\ y_{t-1} \\ R_{t-1} \\ \pi_t \\ y_t \end{bmatrix} +$$

$$\underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ \beta_r \end{bmatrix}}_{\equiv B_1} R_t + \begin{bmatrix} u_{\pi,t+1} \\ u_{y,t+1} \\ u_{R,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

which - using that the policy rule in (A.1) has the form $R_t = -Fx_t$ where $x_t = [x_{1,t}\ x_{2,t}]'$ - can be rewritten as

$$A_0 \begin{bmatrix} \varepsilon_{\pi,t+1} \\ \varepsilon_{y,t+1} \\ \varepsilon_{R,t+1} \\ \pi_t \\ y_t \\ R_t \\ \mathrm{E}_t\pi_{t+1} \\ \mathrm{E}_t y_{t+1} \end{bmatrix} = A_1 \begin{bmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{y,t} \\ \varepsilon_{R,t} \\ \pi_{t-1} \\ y_{t-1} \\ R_{t-1} \\ \pi_t \\ y_t \end{bmatrix} + B_1(-F) \begin{bmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{y,t} \\ \varepsilon_{R,t} \\ \pi_{t-1} \\ y_{t-1} \\ R_{t-1} \\ \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} u_{\pi,t+1} \\ u_{y,t+1} \\ u_{R,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.4}$$

where

$$-F = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & \rho & (1-\rho)\gamma_\pi & (1-\rho)\gamma_y \end{bmatrix}.$$

Finally, we may rewrite (A.4) as

$$\begin{aligned} \begin{bmatrix} x_{1,t+1} \\ \mathrm{E}_t x_{2,t+1} \end{bmatrix} &= A_0^{-1}(A_1 - B_1 F) \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \end{bmatrix} \\ &\equiv A \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \end{bmatrix}. \end{aligned}$$

To solve the model, we then apply the solution algorithm described in Söderlind (1999).

The solution of the model can be written on the form

$$Y_t = \mathbf{C}X_t. \tag{A.5}$$

where $Y_t \equiv [y_t\ \pi_t\ R_t\ \mathrm{E}_t y_{t+1}\ \mathrm{E}_t\pi_{t+1}]'$, $X_t \equiv [y_{t-1}\ \pi_{t-1}\ R_{t-1}\ \varepsilon_{y,t}\ \varepsilon_{\pi,t}\ \varepsilon_{R,t}]'$, and $\mathbf{C}$ is a 5×6 matrix.

Appendix B Estimating the model with FIML

To estimate the model with FIML, we partition the models analytical solution given by (A.5) as

$$Z_t = \mathbf{C}_Z Z_{t-1} + \mathbf{C}_\varepsilon \varepsilon_t \tag{B.6}$$

where $Z_t \equiv [y_t\ \pi_t\ R_t]'$, $\varepsilon_t \equiv [\varepsilon_{y,t}\ \varepsilon_{\pi,t}\ \varepsilon_{R,t}]'$,

$$\mathbf{C}_Z \equiv \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

and

$$\mathbf{C}_\varepsilon \equiv \begin{bmatrix} c_{14} & c_{15} & c_{16} \\ c_{24} & c_{25} & c_{26} \\ c_{34} & c_{35} & c_{36} \end{bmatrix}$$

where c_{ij} denotes the ith row and the jth column in the $\mathbf{C}$ matrix in (A.5). By defining

$$\boldsymbol{\rho} \equiv \begin{bmatrix} \rho_y & 0 & 0 \\ 0 & \rho_\pi & 0 \\ 0 & 0 & \rho_R \end{bmatrix}$$

we may rewrite (A.2) as

$$\varepsilon_t = \boldsymbol{\rho}\varepsilon_{t-1} + u_t \tag{B.7}$$

where $u_t \equiv [u_{y,t}\ u_{\pi,t}\ u_{R,t}]'$ and $u \sim i.i.d.\ N(\mathbf{0}, \mathbf{D})$. By using (B.7) in (B.6), we obtain

$$\begin{aligned} Z_t &= \mathbf{C}_Z Z_{t-1} + \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1} \left(Z_{t-1} - \mathbf{C}_Z Z_{t-2}\right) + \mathbf{C}_\varepsilon u_t \\ &\equiv \left(\mathbf{C}_Z + \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1}\right) Z_{t-1} - \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1} \mathbf{C}_Z Z_{t-2} + v_t. \end{aligned} \tag{B.8}$$

Equation (B.8) forms the basis for the computation of the log-likelihood function. From (B.8), it follows that v_t is distributed as $N(\mathbf{0}, \boldsymbol{\Sigma}_v)$ where $\boldsymbol{\Sigma}_v \equiv \mathbf{C}_\varepsilon \mathbf{D} \mathbf{C}_\varepsilon'$ and that

$$Z_t | Z_{t-1}, Z_{t-2} \sim N\left(\left(\left(\mathbf{C}_Z + \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1}\right) Z_{t-1} - \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1} \mathbf{C}_Z Z_{t-2}\right), \boldsymbol{\Sigma}_v\right). \tag{B.9}$$

Using (B.8) and (B.9), we can then form the conditional log-likelihood function $\ln \ell\left(\mathbf{C}_Z, \mathbf{C}_\varepsilon, \boldsymbol{\rho}\right)$ as (see Hamilton, 1994, Chapter 9)

$$\ln \ell\left(\mathbf{C}_Z, \mathbf{C}_\varepsilon, \boldsymbol{\rho}\right) = -\left(Tn/2\right) \ln\left(2\pi\right) - \left(T/2\right) \ln\left|\boldsymbol{\Sigma}_v\right| - \left(1/2\right) \sum_{t=1}^{T} v_t' \boldsymbol{\Sigma}_v v_t \tag{B.10}$$

where v_t is computed using (B.8). Note that the conditional log-likelihood function in (B.10) does not impose any restrictions on the covariance matrix $\boldsymbol{\Sigma}_v$. Maximization of (B.10) is done with respect to a subset of the parameters in (A.1) and (A.2), i.e. the parameters $\sigma^2_{u_y}$, $\sigma^2_{u_\pi}$ and $\sigma^2_{u_R}$ in $\mathbf{D}$ are not included in the maximization. Instead, they are computed using the FIML estimates as $\hat{\mathbf{D}} = \left(\hat{\mathbf{C}}_\varepsilon\right)^{-1} \hat{\boldsymbol{\Sigma}}_v \left(\hat{\mathbf{C}}_\varepsilon^{-1}\right)'$.

If it is desirable to maximize the likelihood function also with respect to the parameters in $\mathbf{D}$, the covariance restrictions are imposed directly in the conditional log-likelihood function which then becomes

$$\ln \ell\left(\mathbf{C}_Z, \mathbf{C}_\varepsilon, \boldsymbol{\rho}, \mathbf{D}\right) = -\left(Tn/2\right) \ln\left(2\pi\right) - \left(T/2\right) \ln\left(\det\left(\mathbf{C}_\varepsilon^{-1}\right)^2\right) + \left(T/2\right) \ln\left(\det\left(\mathbf{D}\right)\right) - \left(1/2\right) \sum_{t=1}^{T} u_t' \mathbf{D} u_t \tag{B.11}$$

where u_t is computed using (B.8) as

$$u_t = \mathbf{C}_\varepsilon^{-1} \left(Z_t - \left(\mathbf{C}_Z + \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1}\right) Z_{t-1} + \mathbf{C}_\varepsilon \boldsymbol{\rho} \mathbf{C}_\varepsilon^{-1} \mathbf{C}_Z Z_{t-2}\right).$$