Studies in Nonlinear Dynamics & Econometrics

Volume 8, *Issue* 2 2004 *Article* 14

LINEAR AND NONLINEAR DYNAMICS IN TIME SERIES
ESTELLA BEE DAGUM AND TOMMASO PROIETTI, EDITORS

Inference and Forecasting for ARFIMA Models With an Application to US and UK Inflation

Jurgen A. Doornik* Marius Ooms†

Inference and Forecasting for ARFIMA Models With an Application to US and UK Inflation

Jurgen A. Doornik‡ Marius Ooms§

*Nuffield College, Oxford, jurgen.doornik@nuffield.ox.ac.uk
†Free University Amsterdam, mooms@feweb.vu.nl
‡Nuffield College, Oxford, jurgen.doornik@nuffield.ox.ac.uk
§Free University Amsterdam, mooms@feweb.vu.nl

Studies in Nonlinear Dynamics & Econometrics is produced by The Berkeley Electronic Press (bepress). http://www.bepress.com/snde

Inference and Forecasting for ARFIMA Models With an Application to US and UK Inflation*

Jurgen A. Doornik and Marius Ooms

Abstract

Practical aspects of likelihood-based inference and forecasting of series with long memory are considered, based on the arfima(p; d; q) model with deterministic regressors. Sampling characteristics of approximate and exact first-order asymptotic methods are compared. The analysis is extended using modified profile likelihood analysis, which is a higher-order asymptotic method suggested by Cox and Reid (1987). The relevance of the differences between the methods is investigated for models and forecasts of monthly core consumer price inflation in the US and quarterly overall consumer price inflation in the UK.

*Jurgen A. Doornik Nuffield College University of Oxford Oxford OX1 1NF United Kingdom e-mail: jurgen.doornik@nuffield.ox.ac.uk. Marius Ooms Department of Economics Free University of Amsterdam 1081 HV Amsterdam The Netherlands e-mail: mooms@feweb.vu.nl

1 Introduction

Long-memory models have become increasingly popular as a tool to describe economic time series, ever since Mandelbrot suggested to use such models. There is a large literature, much of it focussing on non- and semi-parametric methods. The emphasis here is on modelling and forecasting using methods based on maximum likelihood and regression for the Gaussian fractionally integrated ARMA model (ARFIMA). This allows flexible modelling of the long-run behaviour of the series, and often provides a good description for forecasting.

Useful entry points to the literature are the surveys by Robinson (1994) and Baillie (1996), who consider the developments in the econometric modelling of long memory, and Beran (1992), who reviews long-memory modelling in other areas. The monograph of Beran (1994) discusses most of the central issues, including forecasting.

The Gaussian ARFIMA(p, d, q) model, introduced more carefully in the next section, is specified by the orders of the autoregressive and the moving-average parts of the model, p, and q, as well as the order of differencing, $i + d$. The empirical modelling process consists of three stages: identification, estimation and testing. This may be followed by forecasting.

The first stage determines the integer part i, together with p and q. Our starting point for the second stage is exact maximum likelihood estimation (EML, see §2.1). In contrast to standard ARMA models, the fractional parameter d is estimated, jointly with the AR, MA and regression coefficients. The simple EML estimator of d can be severely biased in the presence of unknown nuisance parameters for regressor variables, even if there is only a constant to measure the unknown mean. Therefore, we also consider the modified profile likelihood method (MPL, §2.2), which incorporates a bias correction. Finally, we estimate by nonlinear least squares (NLS, §2.3), which is easier to implement, and does not require imposition of stationarity.

At the third stage we use diagnostic checks in the form of tests for normality, ARCH effects and neglected serial correlation in the form of the Portmanteau statistic. Specification tests are considered, but, for Wald type, do depend on reliable standard errors. The ARFIMA model enables interesting and easy-to-use tests for the null of short-memory stationarity ($d = 0$), as well as for the null of unit-root nonstationarity ($d = 1$). These tests complement more widely used KPSS and Dickey–Fuller type tests.

Analogous to the distinction between NLS and EML, there are two ways to compute the forecasts, differing in the treatment of pre-sample values. The optimal methods we employ are explicitly based on finite sample information sets; we also use 'naive' methods which ignore this issue.

The aims we have are as follows:

- illustrate the feasibility of EML and MPL estimation for ARFIMA models;
- investigate whether the three estimation methods, EML, MPL and NLS, give empirically similar results;

- compare the optimal and naive methods for forecasting.

These issues are studied with reference to consumer price inflation for the US and UK. Inflation series display long non-periodic cycles typical of long-memory time series. This leads to significantly positive estimates of the order of integration which determines the rate of growth of out-of-sample forecast intervals of the log of the corresponding price level. So, d is often the primary parameter of interest and of crucial importance for studies of long-run variation of e.g. indexed outlays.

The organization of this paper is as follows. The next section introduces the ARFIMA model, and the estimators that are studied. Section §3 estimates a model for quarterly UK inflation, and presents forecasts of the log-price level. The next section then studies this model through a Monte Carlo analysis. Section §5 estimates a model for monthly inflation for the US, which is followed by a parametric bootstrap analysis. Finally, in §6 we consider the impact of neglected GARCH-structure for the error term.

2 Estimation and forecasting for the arfima model

The ARFIMA(p, d, q) model for y_t is written as

$$\Phi(L)(1-L)^d(y_t-\mu_t)=\Theta(L)\varepsilon_t, \quad t=1,\ldots,T. \tag{2.1}$$

where $\Phi(L)=(1-\phi_1 L-\ldots-\phi_p L^p)$ is the autoregressive polynomial and $\Theta(L)=(1+\theta_1 L+\ldots+\theta_q L^q)$ is the moving average polynomial in the lag operator L; p and q are integers, d is real. $(1-L)^d$ is the fractional difference operator defined by the following binomial expansion:

$$(1-L)^d=\sum_{j=0}^{\infty}\delta_j L^j=\sum_{j=0}^{\infty}\left(\begin{array}{c} d \\ j \end{array}\right)(-L)^j.$$

We assume $\varepsilon_t \sim \mathsf{NID}(0,\sigma_\varepsilon^2)$, and write μ_t for the mean of y_t. The ARMA-part of the model is assumed invertible and stationary. In addition, $\Theta(z)=0$ and $\Phi(z)=0$ do not have common roots. In that case $z_t=y_t-\mu_t$ is integrated of order d, denoted $I(d)$.

The properties of z_t depend crucially on the value of d:

it is covariance stationary if $d<0.5$, with long memory if $d>0$ (see Hosking (1981)). When $0<d<0.5$ the autocovariance function exhibits hyperbolic decay: $\gamma_k \sim ck^{2d-1}$ for $k\to\infty$, where c denotes a finite nonzero constant. The spectral density $f_z(\omega)$ near zero is also hyperbolic: $\lim_{\omega\to 0} f_z(\omega)\omega^{2d}$ exists and is finite. For $-0.5<d<0$ the process is called intermediate memory or 'overdifferenced'. In that case the inverse autocorrelations decay hyperbolically.

Odaki (1993) used the following condition for invertibility of z_t: convergence of the mean squared error of the AR(∞)-based one-step-ahead prediction, $\mathsf{MSE}(\widehat{z}_{t|T})$, to the innovation variance σ_ε^2 as $T\to\infty$. The AR(∞) representation of z_t is defined as

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + \varepsilon_t. \tag{2.2}$$

In obvious notation: $\Pi(L) = \sum_{j=0}^{\infty} \pi_j L^j = \Theta(L)^{-1}\Phi(L)(1-L)^d$ and $\pi_0 = 1$. Note that there is an AR unit root for $d > 0$. When pre-sample values, i.e. z_j for $j < 0$, are set to zero in forecasting, we call the corresponding predictions 'naive' forecasts. These predictions are optimal if the observations are known into the infinite past. The corresponding one-step-ahead forecast errors are labelled naive residuals, denoted by $\tilde{e_t}$. Forecasting is discussed in more detail in §2.5. The coefficients of $(1-L)^d$, δ_j, are easily computed: $\delta_0 = 1$ and

$$\delta_j = \prod_{0<k\leq j} \frac{k-1-d}{k}, \quad j = 1, 2, \ldots$$

The MA representation of z_t:

$$z_t = 1 + \sum_{j=1}^{\infty} \psi_j \varepsilon_t = \Psi(L)\varepsilon_t = \Phi(L)^{-1}(1-L)^{-d}\Theta(L)\varepsilon_t, \tag{2.3}$$

has an MA-unit root when $-1 < d \leq -0.5$. Note that $\psi_j \to 0$, when $j \to \infty$ for $d < 1$. The process is therefore mean-reverting in this case, and innovations ε_t only have a transitory effect on the time-series process. In fact: $\psi_k \sim ck^{d-1}$ for $-0.5 < d < 1$.

The remainder of this section reviews EML, MPL and NLS estimation. These are all implemented in the ARFIMA-package of Doornik and Ooms (1999), which is a class of procedures in the programming language Ox, see Doornik (2001). The package also implements the forecasting methods discussed below.

2.1 Exact maximum likelihood (EML)

Based on the normality assumption and with a procedure to compute the autocovariances in the $T \times T$ covariance matrix $\mathbf{\Sigma} = \sigma_\varepsilon^2 \mathbf{R}$ of a $T \times 1$ vector of observations $\mathbf{y}$, the log-likelihood for the ARFIMA(p, d, q) model (2.1) with k regressors is

$$\log L\left(d, \phi, \theta, \beta, \sigma_\varepsilon^2\right) = -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log\sigma_\varepsilon^2 - \frac{1}{2}\log|\mathbf{R}| - \frac{1}{2\sigma_\varepsilon^2}\mathbf{z}'\mathbf{R}^{-1}\mathbf{z}, \tag{2.4}$$

where $\mathbf{z} = \mathbf{y} - \mathbf{X}\beta$. When σ_ε^2 and β are concentrated out, the resulting normal profile likelihood function becomes:

$$\log L_P\left(d, \phi, \theta\right) = c - \frac{1}{2}\log|\mathbf{R}| - \frac{T}{2}\log\widehat{\mathbf{z}}'\mathbf{R}^{-1}\widehat{\mathbf{z}}, \tag{2.5}$$

where $\widehat{\mathbf{z}} = \mathbf{y} - \mathbf{X}\widehat{\beta}$ and

$$\widehat{\beta} = (\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}. \tag{2.6}$$

The core of the EML method is the computation of the autocovariances as a function of the parameters of a stationary ARFIMA model. Hosking (1981) presents an effective method to compute the ACF for an ARFIMA$(1, d, 1)$ process. This is extended by Sowell (1987) to the general case, using recursive evaluation of the hypergeometric functions that are involved. Doornik and Ooms (2003) give a small improvement to this method which enhances numerical stability. They also show that the computational cost of Sowell's method for the ACF is negligeable compared to the likelihood evaluation (2.5), which contains the inverse and determinant of a $T \times T$ covariance matrix. Doornik and Ooms (2003) show how estimation can be achieved efficiently using storage of order T, and computational effort of order T^2. As a consequence EML is sufficiently fast to be used in samples of up to several thousand observations and bootstrap methods.

2.2 Modified profile likelihood (MPL)

A good estimate of d is required for estimating the variance of the sample mean, especially for inference about the mean of y_t, and for forecasting. In most cases, the unknown mean μ_t is a function of nuisance parameters, whose presence has an adverse effect on the finite sample behaviour of the standard maximum likelihood estimator of d. When μ_t is estimated, either by simple regression, or jointly with d by maximizing the profile likelihood, $\widehat{d}$ can be severely biased, even in the simplest ARFIMA$(0, d, 0)$ model. Smith Jr et al. (1997) suggest to overdifference the data to remove the constant term, and then directly estimate $d - 1$ essentially without bias. Our results show that this procedure is also effective when there are additional regressors.

The modified profile likelihood is a concept from higher order asymptotic theory, see Cox and Reid (1987). The aim is to develop more accurate inference on parameters of interest in the presence of (a large number) of nuisance parameters, see (Barndorff-Nielsen and Cox, 1994, Ch. 4) for a motivation and examples. An and Bloomfield (1993) derive the modified profile likelihood, $\log L_M$, for the regression model with stationary ARFIMA-errors:

$$\log L_M\left(d,\phi,\theta\right) = c + \left(\frac{1}{T}-\frac{1}{2}\right)\log|\mathbf{R}| - \frac{1}{2}\log\left|\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\right| - \left(\frac{T-k-2}{2}\right)\log\widehat{\mathbf{z}}'\mathbf{R}^{-1}\widehat{\mathbf{z}}. \tag{2.7}$$

They show that the expectation of the score of (2.7) for (d, ϕ, θ) at the true parameter values is $O(T^{-1})$, whereas this expectation is $O(1)$ for the score of (2.5). This higher order bias correction provides the main argument for the better behaviour of the MPL estimator over the EML estimator. The EML and MPL estimators require $\mathbf{\Sigma}$ and its inverse to exist, and therefore require $d < 0.5$.[1] The assumption of normality plays an important role in the derivation of the optimality of both estimators.

[1] The restriction $-1 < d <= 0.49999$ is imposed in EML and MPL estimation, together with roots of the AR polynomial between -0.9999 and 0.9999.

The Monte Carlo experiments of An and Bloomfield (1993), and the results in Hauser (1999), show that, for simple ARFIMA$(1, d, 1)$ models, MPL reduces bias for $\widehat{d}$, and that it leads to more accurate inference in finite samples. Hauser (1999) uses the OLS-estimate of β in the MPL estimator in order to reduce the number of computations; we use the appropriate GLS estimator (2.6).

2.3 Nonlinear least squares (NLS)

Beran (1995) develops an approximate maximum likelihood estimator based on minimizing the sum of squared naive residuals, which is also applicable for nonstationary ARFIMA-processes with $d > 0.5$. The approximate log likelihood is:

$$\log L_A\left(d, \phi, \theta, \beta\right) = c - \frac{1}{2} \log \frac{1}{T-k} \sum_{t=2}^{T} \tilde{e}_t^2, \tag{2.8}$$

where $\tilde{e}_t$ are the one-step-prediction errors from the naive predictions defined near (2.2). Beran proves asymptotic efficiency and normality of the resulting estimators for (d, ϕ, θ). Beveridge and Oickle (1993) and Chung and Baillie (1993) present Monte Carlo evidence which suggest it to be a good estimator for ARFIMA(0,d,0) models with unknown mean. Chung and Baillie (1993) called this estimator the conditional sum-of-squares estimator. We call it the nonlinear least-squares estimator.

2.4 Inference

The asymptotic efficiency and normality of the maximum likelihood estimators for (d, ϕ, θ) and β have been established by Dahlhaus (1989) and Dahlhaus (1995).

For all estimation methods, we use the numerical Hessian as the basis for the estimated covariance matrix of the parameters. The covariance matrix of $\widehat{\beta}$ is of the familiar GLS type:

$$\mathsf{var}\widehat{\beta} = \widehat{\sigma}_\varepsilon^2 (\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-1}.$$

To our knowledge, an MPL estimator for $\widehat{\sigma}_\varepsilon^2$, has not been derived, except for special cases such as the standard linear model without dynamics, where it equals the familiar unbiased OLS-estimator: $\widehat{\sigma}_\varepsilon^2 = (T-k)^{-1}\mathbf{e}'\mathbf{e}$, see e.g. (Barndorff-Nielsen and Cox, 1994, example 4.9). For "MPL-inference" on β we employ this OLS-formula for $\widehat{\sigma}_\varepsilon^2$.[2]

2.5 Forecasting

The residuals e_t of EML and $\tilde{e}_t$ of NLS estimation are the results of two different methods of prediction: best linear unbiased prediction and 'naive' prediction, which can also be applied out-of-sample.

[2] Kiviet and Phillips (1998) discuss developments in bias corrections of $\widehat{\sigma}_\varepsilon^2$ in AR-models and suggest increasing the degrees-of-freedom correction with the number of estimated AR parameters. The correction of Lieberman (2001) is one higher, owing to the estimation of d.

The best linear prediction of z_{T+H}, given the information in $\mathbf{z}$ and knowing the parameters of the ARFIMA process, is given by

$$\widehat{z}_{T+H|T} = (\gamma_{T-1+H} \cdots \gamma_H)\,(\boldsymbol{\Sigma}_T)^{-1}\mathbf{z} = \mathbf{q}'_H\mathbf{z} \tag{2.9}$$

see e.g. (Beran, 1994, §8.7) or (Brockwell and Davis, 1993, §5.1). Again, this can be viewed as a regression of z_{T+H} on $\mathbf{z}$, and $(\boldsymbol{\Sigma})_T^{-1}\mathbf{z}$ can be computed efficiently using a Durbin–Levinson type algorithm. Let $\mathbf{Q}_{H|T}\mathbf{z}$ denote the optimal forecast for $\widehat{\mathbf{z}}_{H|T} = (\widehat{z}_{T+1|T} \cdots \widehat{z}_{T+H|T})'$, then $\mathsf{var}(\widehat{\mathbf{z}}_{H|T} - \mathbf{z}_H) = \boldsymbol{\Omega}_{H|T} = \boldsymbol{\Sigma}_H - \mathbf{Q}_{H|T}\boldsymbol{\Sigma}_T\mathbf{Q}'_{H|T}$. The diagonal elements give the mean squared errors of the optimal j-step ahead forecasts $j = 1, \ldots, H$. It is often of interest to forecast partial sums of z_t, e.g. when log price-level predictions are constructed as partial sums of inflation forecasts. The variance matrix of cumulative prediction errors is then easily computed as $\mathbf{C}\boldsymbol{\Omega}_{H|T}\mathbf{C}'$ where $\mathbf{C}$ is a lower triangular matrix of only ones.

The naive method recursively predicts $\tilde{z}_{T+1}, \tilde{z}_{T+2}, \ldots$ using AR-representation (2.2) up to order $T, T+1, \ldots$. In that case, pre-sample values are set to zero, and the predictions are optimal if the observations are known into the infinite past. Corresponding variances of $\tilde{z}_{T+H}$ are computed using the MA-coefficients of (2.3):

$$\mathsf{var}(\tilde{z}_{T+H}) = \sigma_\varepsilon^2 \left(1 + \sum_{j=1}^{H-1} \psi_j^2\right). \tag{2.10}$$

Again, $\tilde{z}_{T+H|T}$ converges to $\widehat{z}_{T+H|T}$ as $T \to \infty$.

2.6 Monte Carlo simulation and bootstrap inference

All that is required for simulation, in addition to efficient implementations of estimators and tests, are exact drawings from the DGP of an ARFIMA process. Let $\mathbf{PP}'$ be a Choleski factorization of $\boldsymbol{\Sigma}$ then drawings $\mathbf{y}$ are conveniently generated as $\mathbf{y} = \mathbf{P}\varepsilon + \mathbf{X}\beta$ where ε is a vector of independent standard normal drawings. For large samples sizes, storage of the triangular matrix $\mathbf{P}$ may be problematic. In that case, an inverse version of Durbin's algorithm can be applied, see Doornik and Ooms (2003).

Monte Carlo experiments are used to compare the estimates from the three estimators when applied to the same data generation process, based on findings from the UK model. The parametric bootstrap, which in our implementation amounts to a Monte Carlo experiment where the parameters are given by the empirical model, is used to assess the US results. Unlike a Monte Carlo experiment, each model is simulated at its empirical estimates, instead of using the same set of DGP values. This will enable us to explore the impact on inference and forecasting of the different baseline results.

Bootstrap samples can also be used for parametric bootstrap tests: the time series is resampled using the estimates of parameters under a null hypothesis of interest. Compute the test statistics for the observed sample, as well as for each bootstrap sample. If the estimated bootstrap p-value, i.e. the proportion of simulated test statistics that exceeds the observed test statistic, is smaller than our significance level, the null hypothesis is rejected. In many circumstances one may expect bootstrap inference to be more accurate than standard asymptotic inference, see Davidson and MacKinnon (1999b).

3 Modelling Inflation in the UK

The UK series is the quarterly 'all items retail price index'[3] for the period 1959.1–2002.2. We construct the corresponding inflation series by defining $p_t = \Delta \log P_t$, and multiplying by 400 to obtain annual percentages. The series exhibits long memory and clear seasonality, see Figure 1a.

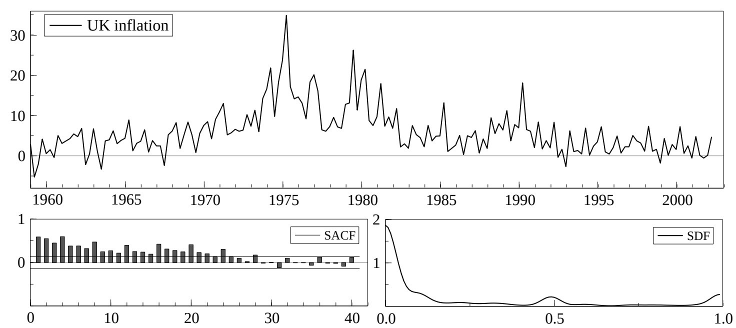

Figure 1: UK quarterly inflation rates, with sample autocorrelation function and spectral density.

The time-series plot shows slow mean reversion of inflation. There are two pronounced peaks in inflation in the 1970s: in the second quarter of 1975, and the third quarter of 1979. These could be traced back to the impact of the first and second oil-price shocks and VAT adjustments, and will require dummy variables in our model. The 1975 peak was a culmination of oil-price shocks and domestic factors (see the discussion in Hendry (2001) who builds a structural model for UK inflation, extending over more than a century.) Early 1979 was the so-called 'winter of discontent', followed by the election of Margaret Thatcher in May. The first budget in June introduced a switch from direct to indirect taxation, with a sharp increase in VAT. A few weeks later this was followed by the second oil price shock. These effects caused a sharp increase in measured inflation. We capture both effect in one variable called VAT, which is unity for 1975.2 and 1979.3.

Figure 1b contains the corresponding sample autocorrelation function (SACF) $\widehat{\gamma}_k/\widehat{\gamma}_0$, for lags of 1 up to 40 quarters. The decay in the SACF is very slow, indicating a d close to, or inside the nonstationary region $[0.5, \infty)$. There is also

[3]This is the headline RPI, with 1987=100. It is available as the series CHAW from the UK Office of National Statistics, www.statistics.gov.uk. Prior to 1987, the RPI is extended backward using the series CZBH.

clear evidence of the seasonality in the SACF. Some experimentation leads us to a very simple form to capture the seasonality: the up-swing in the second quarter is off-set in the third quarter, so we use the variable $Q_2 - Q_3$. The non-parametric estimate of the spectral density function (SDF) in Figure 1c shows a clear peak near frequency zero.

3.1 Inference

An adequate, yet parsimonious, approximation to the second-order characteristics for UK inflation is given by the following ARFIMA model with zero-frequency long-memory and short-memory seasonality:

$$(1 - \phi_p L^p)(1 - L)^d (y_t - x_t' \beta) = (1 + \theta_q L^q)\varepsilon_t,$$

where $p = q = 4$ for the quarterly UK data, and x_t consists of a constant and dummy variables (VAT and $Q_2 - Q_3$). Table 1 presents estimation results for the three methods EML, MPL and NLS.

The UK estimates for d vary from 0.47 to 0.59. The results in Table 1 illustrate an important practical problem of estimating a d near 0.5: EML delivers an estimate smaller than 0.5, whereas MPL did not converge. Therefore we re-estimate after differencing the data, based on $(1-L)^d = (1-L)^{d-1}(1-L)$:

$$(1 - \phi_4 L^4)(1 - L)^{d_\Delta}(\Delta y_t - \Delta x_t' \beta) = (1 + \theta_4 L^4)\varepsilon_t, \tag{3.1}$$

using ΔVAT and $\Delta(Q_2 - Q_3)$ in the mean. The tabulated value is the MPL estimate of d which is $1 + \widehat{d}_\Delta$. As a consequence of differencing we can no longer identify a value for the mean of inflation: the constant drops out of the regressor set. Note that MPL and EML differ only if there are regressors in the model, so except for the small effect of the dummy variables, the MPL results can also be viewed as EML estimates obtained by modelling the first differences. NLS does not require such first differencing to restrict $d < 0.5$. There is, however, a change in the interpretation of the constant term. For $-0.5 < d < 0.5$ the constant term represents the mean of inflation, but for $0.5 < d < 1$ it should be interpreted as the mean growth rate. The constant is of course unidentified (cannot be estimated) if $d = 1$. In the UK case the mean growth estimate is apparently not well identified in the data. The UK inflation seems nonstationary, but one cannot easily reject $d = 0.5$ for the UK data.

The last four lines of Table 1 report test residual diagnostics for the model, as reported by PcGive (Hendry and Doornik (2001)), and applied to the EML and MPL residuals e_t and to the naive NLS residuals $\tilde{e}_t$. The normality test is that of Doornik and Hansen (1994), Portmanteau is the Ljung and Box (1978) statistic with the used lag length in parentheses, ARCH(4) is the LM test for fourth order ARCH effects, Engle (1982). P-values of the statistics are reported between square brackets. There is no clear evidence against the white noise assumption but some ARCH effects are detected. This corresponds to the higher volatility during the period of high inflation. The results for the EML-e_t, MPL-e_t and $\tilde{e}_t$ are qualitatively similar.

Table 1: Estimation results for UK inflation, 1959.1–2002.2 ($T = 174$)

	EML	MPL	NLS
dependent variable	$400p_t$	$400\Delta p_t$	$400p_t$
$\widehat{d}$	0.473 (0.032)	0.563 (0.069)	0.590 (0.076)
$\widehat{\phi}_4$	0.872 (0.079)	0.918 (0.063)	0.612 (0.18)
$\widehat{\theta}_4$	−0.695 (0.12)	−0.742 (0.095)	−0.434 (0.20)
Constant	3.60 (15)	—	1.80 (9.7)
VAT	15.6 (1.7)	15.2 (1.7)	15.0 (1.6)
$Q_2 - Q_3$	2.58 (0.51)	2.45 (0.61)	3.00 (0.33)
$\widehat{\sigma}$	2.836	2.856	2.745
Normality	2.02 [0.36]	3.43 [0.18]	2.92 [0.23]
Portmanteau(13)	12.1 [0.28]	10.4 [0.41]	10.3 [0.41]
Portmanteau(25)	26.6 [0.23]	23.8 [0.36]	28.4 [0.16]
ARCH(4)	4.21 [0.003]	4.92 [0.001]	4.01 [0.004]

3.2 Forecasting

An important goal of long-memory time-series modelling is to perform inference on long-range forecasts. How do the differences in estimates of d and β translate into the location and width of forecast intervals for inflation and the log price level? We present the main picture in Figure 2 for the UK, which displays forecast intervals up to a horizon of 8 years together with nominal 95% confidence intervals. The dependent variable was 'seasonally adjusted', by taking $p_t - 2.5(Q_2 - Q_3)/400$, using actual growth rates instead of annual percentages. The estimation sample was reduced by one observation to 2002.1 to avoid the seasonal effect in the end-of sample observation when re-integrating.

The different panels clearly show how the rate of growth of the forecast interval depends on $\widehat{d}$. The effective orders of integration are about 0.47, -0.44 and 0.59 for the top graphs and 1.47, 1.56 and 1.59 for the bottom graphs.

The EML estimate of 0.47 indicates stationarity for the level of UK inflation. With a d so close to 0.5 we observe that the forecast and its variance converge only very slowly to the unconditional mean and variance. The (short-memory) deviations in the stationary seasonal pattern are apparently predictable. This shows most clearly in the forecasts for the changes in inflation ($d = -0.44$) which form the basis of the MPL projections. Here the forecasts converge much faster to the unconditional (zero) mean and variance of the process.

The NLS forecasts for inflation are based on $d = 0.59$, which is clearly above 0.5. The regression constant in the model now determines a (downward) trend inflation. The variance of the H-step ahead forecasts grows proportionally to H^{2d-1} for nonstationary $d > 0.5$, assuming $H/T \to 0$, see (Beran, 1994, §8.6). Therefore, the width of the NLS forecast interval for inflation is proportional to the corresponding square root: $cH^{0.59-0.5}$.

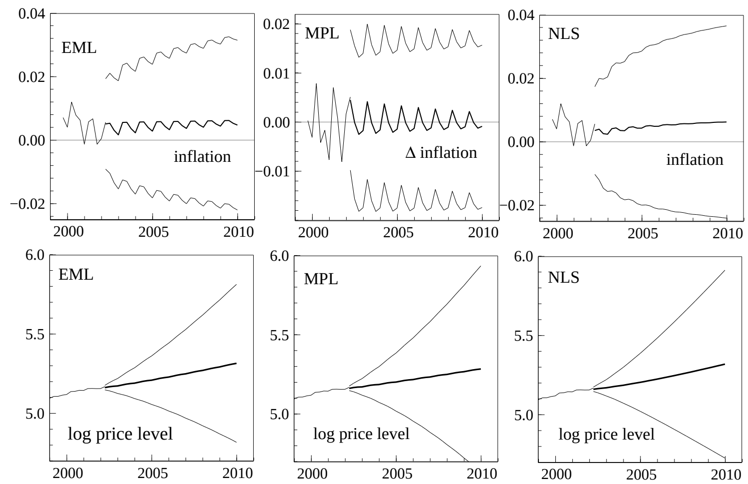

Figure 2: Forecasts of (Δ) inflation levels and log price levels with 95% confidence bands for the EML, MPL and NLS estimates for the UK model. Optimal forecasts for EML and MPL, naive forecasts for NLS.

The different forecasts for the log price level are displayed in the bottom graphs of Figure 2. All these forecasts show clearly that the UK price level is not mean reverting: $d \gg 1$. The EML price level forecast is linearly trending and the width of the forecast interval is proportional to $H^{0.97}$. Therefore, the interval grows (nearly) at a linear rate as well. The forecast interval is eventually bounded from below, although this does not seem important for relevant forecast horizons. The MPL price-level forecasts are also linearly trending, since we have integrated the zero mean Δinflation series twice. The slope resembles the slope in the EML forecasts. The forecast interval grows proportionally to $H^{1.06}$. The NLS forecast approaches a quadratic trend and the growth of the forecast interval reflects the covariance nonstationarity most clearly.

Table 2: Forecast results for UK log-price level 1 to 32 quarters ahead

		EML		MPL				NLS	
		optimal: $\widehat{z}_{T+H}$		optimal: $\widehat{z}_{T+H}$		naive: $\tilde{z}_{T+H}$		naive: $\tilde{z}_{T+H}$	
H	Date	forecasts	RMSE	forec.	RMSE	forec.	RMSE	forec.	RMSE
1	2002.2	5.1637	(0.0071)	5.1636	(0.0071)	5.1637	(0.0071)	5.1622	(0.0069)
2	2002.3	5.1691	(0.0127)	5.1688	(0.0133)	5.1690	(0.0133)	5.1663	(0.0129)
4	2003.1	5.1741	(0.0236)	5.1726	(0.0260)	5.1731	(0.0259)	5.1714	(0.0255)
12	2005.1	5.2104	(0.0764)	5.2031	(0.0915)	5.2060	(0.0911)	5.2059	(0.0912)
24	2008.1	5.2718	(0.174)	5.2515	(0.221)	5.2613	(0.220)	5.2713	(0.210)
32	2010.1	5.3167	(0.249)	5.2857	(0.325)	5.3021	(0.322)	5.3209	(0.296)

Sample: 1959(1) - 2002(1). See Table 1 for model and parameter estimates.

Table 2 provides numerical evidence corresponding to Figure 2; it also contrasts the optimal and naive forecasts. The former involve formulae for projection into the finite past. The weight of past observations declines only slowly in long-memory models and one can expect a substantial difference with naive forecasts which use projection as if the infinite past is known. However, Table 2 shows that the differences are not that big. The results for MPL required us to integrate the forecasts for Δinflation twice. Evenso, the log price-level forecasts 8-years ahead differ only by .016, with estimated standard error of 0.325 and 0.322. In sum, we see that the choice of estimator matters much more than the choice of predictor. Naturally, the difference between optimal and naive forecasts becomes more pronounced as the sample size decreases.

4 Monte Carlo analysis

It was already noted that the estimate of d close to 0.5 is downward biased for EML. Because the results of the previous section suggest that d is close, and perhaps larger than 0.5, we investigate this using a Monte Carlo experiment. The DGP is chosen to resemble the UK estimates, selecting $d_0 = 0.45$ initially:

$$(1 - 0.9L^4)(1 - L)^{d_0}\left[y_t - 3.5 - 15\text{VAT} - 2.5(Q_2 - Q_3)\right] = (1 - 0.7L^4)\varepsilon_t, \quad (4.1)$$

with $\varepsilon_t \sim N(0, 8)$ and $t = 1, \ldots, 174$. The Monte Carlo uses $M = 10\,000$ replications.

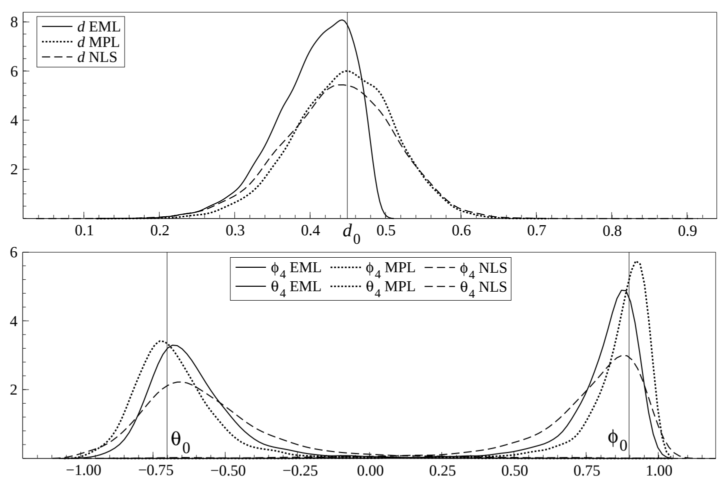

Figure 3: Estimates for $\widehat{d}$ (top panel) and $\widehat{\phi}_4, \theta\phi_4$ (bottom panel) for DGP (4.1), $M = 10,000$.

The solid line in the first panel of Figure 3 shows the distribution of the EML estimates of $\widehat{d}$. The bias towards stationarity is clearly visible. Just one experiment failed to converge, and all others had $\widehat{d} < 0.49999$, so this bias is not an artefact of rejecting non-stationary experiments. The bottom panel gives the distributions for the estimated ARMA parameters.

The dotted line in the first panel of Figure 3 shows the results for the MPL estimates. Here we automatically switched to estimation in first differences when there is a convergence failure with $d > 0.49$. This enables us to estimate a non-stationary model, thus matching our modelling approach for UK inflation. The cost is the loss of identification of the mean. The MPL estimates for $\widehat{d}$ are essentially unbiased, and approximately normally distributed. The NLS estimates of d are quite close to those of MPL, but it has much more problems identifying the short-run dynamics in the form of the ARMA parameters (this difference with EML/MPL was also noticeable from the forecast plots, Fig. 2).

Table 3: Monte Carlo estimates for DGP (4.1) with $d_0 = 0.45$, $M = 10,000$

	EML		MPL		NLS	
	bias	RMSE	bias	RMSE	bias	RMSE
$\widehat{d}$	−0.045	0.07	−0.005	0.07	−0.011	0.08
$\widehat{\phi}_4$	−0.091	0.18	−0.038	0.14	−0.155	0.29
$\widehat{\theta}_4$	0.093	0.20	0.035	0.17	0.148	0.30
Constant	−0.131	12.11	−1.160	10.49	−0.039	13.10
VAT	0.009	1.74	0.006	1.73	0.010	1.75
$Q_2 - Q_3$	−0.001	0.63	−0.001	0.63	0.002	4.45
$\widehat{\sigma}^2$	−0.268	0.88	−0.165	0.87	−0.313	0.92

Table 3 confirms the graphical results. With a DGP that is close to non-stationarity, MPL provides the best estimate of d. In addition, it has a lower bias on the ARMA coefficients, with a lower root mean-square error (RMSE). The estimate of the constant is more biased, because it was set to zero whenever the model was using differenced data.[4]

The scatterplots in the first row of Figure 4 show that d is well identified: $\widehat{d}$ is hardly correlated with $\widehat{\phi_4}$. This is an important check: empirical identification is a necessary condition for the successful application of standard asymptotic results in practice. The scatterplot of $\widehat{\phi}_4$ versus $\widehat{\theta}_4$ shows that both parameters are not well identified individually. However, it is also clear that $\phi_4 + \theta_4 = 0$, i.e. cancelling roots and no seasonal parameters, can be rejected. The line $\phi_4 + \theta_4 = 0$ in the scatter plot indicates this. Note that our EML and MPL implementations restrict ϕ_4 to an interval $[-1 + \delta, 1 - \delta]$, where we chose $\delta = 0.001$; NLS does not impose this restriction.

The sum of the ARMA parameters, $\phi_4 + \theta_4$, can be interpreted as the impulse response of the short-memory component of inflation after one year. This statistic has a much better behaved distribution, see Figure 5, with EML and MPL

[4] We also run the MPL experiment without automatic differencing. In that case, 14 000 experiments are needed to successfully estimate 10 000 stationary models. Of course, this selection causes a bias in d, which at −0.035 is still slightly lower than that of EML.

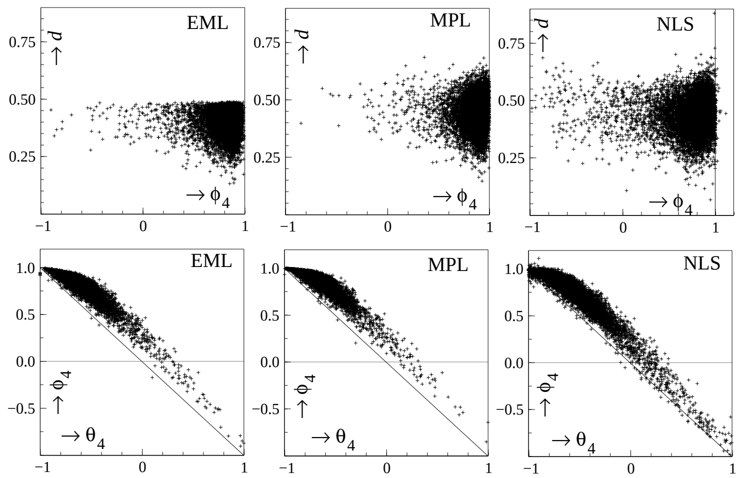

Figure 4: Cross plots of Monte Carlo replications of three estimators for DGP (4.1) with $d_0 = 0.45,\ M = 10,000$.

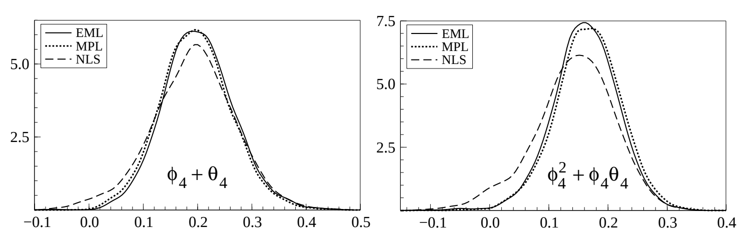

Figure 5: Monte Carlo estimates of impulse responses for short-memory part of DGP (4.1) with $d_0 = 0.45$.

very similar. The short-memory impulse response after 2 years, $\phi_4(\phi_4 + \theta_4)$, is also well behaved, although NLS is again much less reliable.

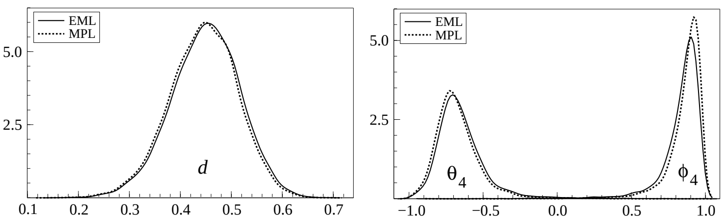

Figure 6: EML estimates of first differences compared to MPL estimates, for DGP (4.1) with $d_0 = 0.45$.

Figure 6 shows that overdifferencing does indeed remove the bias for EML. All EML estimates are now of the form (3.1), while MPL only differences if required (which is in about 30% of estimates). Now both are very similar, although the bias in the ARMA parameters of EML is still somewhat higher. MPL has the advantage that, for d close to 0.5, it is possible to estimate all mean parameters without introducing a bias in d. The overdifferenced EML estimate of d for UK inflation (Table 1) is 0.57 which is compatible with the MPL estimate. Indeed, when we rerun the Monte Carlo experiments with $d_0 = 0.55$ in the DGP (i.e. setting $d_0 = -0.45$ and then integrating once), the EML estimates are near 0.45, while the MPL estimates are again unbiased.

5 Modelling Inflation in the US

The US series is the monthly core inflation in the US, taken as the urban consumer price index for all items except food and energy.[5] The sample period is 1957.1–2003.4. The corresponding inflation series is $100\Delta \log P_t$ to obtain annual percentages. Like UK inflation, the US series exhibits long memory and clear seasonality, see Figure 7. In addition, the effect of rounding in the data is visible prior to 1975. There is a pronounced outlier in July 1980: we will treat this sudden drop in the price level when the US was locked in high inflation as a measurement error.

Our model specification follows Bos et al. (2002) in that we allow for a separate mean of inflation in the period from 1973.7 to 1982.6, as well as a shift in seasonality from 1984 onwards. Finally, we add one additional model-based outlier correction for 1981.10. Table 4 reports the results, omitting the 22 coefficients for the centered seasonals.

The US estimates of $\widehat{d}$ are all close to 0.32, with a standard error of 0.03. This puts it in the stationary region, but with clear evidence of long-memory. The other coefficient estimates also show little difference between the EML and

[5] This series has 1982-84=100, and is available as the series CUUR0000SA0L1E from the US Bureau of Labor Statistics, www.bls.gov.

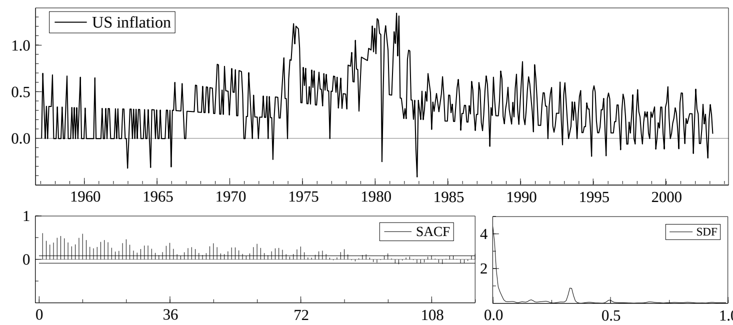

Figure 7: US monthly inflation rates, with sample autocorrelation function and spectral density.

Table 4: Estimation results for US inflation, 1957.11–2003.4 ($T = 546$; centered seasonal estimates omitted)

	EML	MPL	NLS
dependent variable	$100p_t$	$100p_t$	$100p_t$
$\widehat{d}$	0.322 (0.032)	0.318 (0.034)	0.332 (0.032)
$\widehat{\phi}_{12}$	−0.156 (0.33)	0.348 (0.51)	0.641 (0.11)
$\widehat{\theta}_{12}$	0.273 (0.32)	−0.184 (0.52)	−0.633 (0.13)
Constant	0.250 (0.068)	0.249 (0.076)	0.317 (0.065)
73M7–82M6	0.404 (0.056)	0.405 (0.060)	0.387 (0.055)
D80M7	−1.19 (0.16)	−1.20 (0.16)	−1.11 (0.16)
D80M10	−0.543 (0.16)	−0.539 (0.16)	−0.484 (0.16)
$\widehat{\sigma}$	0.167	0.172	0.167
Normality	2.42 [0.30]	3.65 [0.16]	5.23 [0.07]
Portmanteau(36)	40.8 [0.17]	44.1 [0.09]	45.9 [0.07]
Portmanteau(72)	76.0 [0.26]	77.8 [0.22]	89.9 [0.05]
ARCH(12)	6.18 [0.000]	6.27 [0.000]	5.10 [0.000]

MPL estimates. NLS differs mainly in the estimates of the ARMA parameters, and their standard errors. The test outcomes are also somewhat less satisfactory (but all show very significant ARCH effects).

5.1 Bootstrap inference

We use a parametric bootstrap to check the reliability of the estimates. The focus is on the ARMA parameters and intercept, where the main difference is found. In each case, the null hypothesis is the respective model estimate (so EML is simulated using a DGP based on the EML estimates of Table 4, MPL using the MPL estimates, etc.). Only 100 replications are used to illustrate that this can be a useful part of empirical modelling (except for NLS, this takes less than 10 minutes on a 1.6 Mhz Pentium IV).

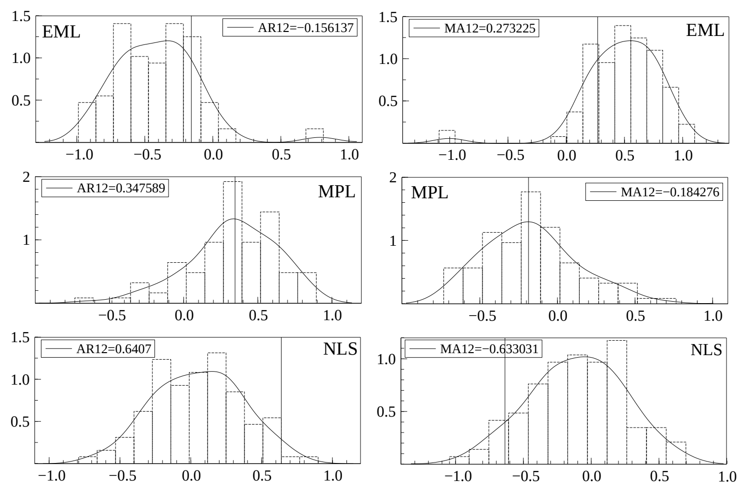

Figure 8: Estimates of parametric bootstrap densities for $\widehat{\phi}_{12}$ and $\widehat{\theta}_{12}$ for the US model of Table 4. See also Table 5.

Figure 8 shows the distribution for the ARMA coefficients of the three models. Even at 100 replications there is a clear difference: the EML estimates are quite biased, while the MPL estimates appear well-behaved. NLS is problematic, because the model has essentially cancelling roots. Estimation takes nearly ten times as long as for the other estimators (and there are 33 convergence failures to get 100 converged models). The NLS estimates are centered on zero, instead of the DGP values. It is quite surprising that with nearly 550 observations there is still such a large difference between the estimators.

For the other parameters in the model, EML, MPL and NLS all have a small bias. However, inference on the intercept and intercept shift is more reliable for MPL than both EML and NLS, see Table 5 (although the number of replications is very small here).

Table 5: Empirical rejection probabilities at 10% and 5% nominal level for two-sided t-tests. US core inflation

	EML		MPL		NLS	
	10%	5%	10%	5%	10%	5%
$t-\widehat{d}$	0.05	0.01	0.03	0.01	0.17	0.10
$t-\widehat{\phi}_{12}$	0.36	0.33	0.18	0.11	0.71	0.63
$t-\widehat{\theta}_{12}$	0.38	0.33	0.19	0.15	0.62	0.48
$t-$Constant	0.18	0.10	0.09	0.05	0.25	0.19
$t-$D73M7-82M6	0.20	0.16	0.17	0.11	0.23	0.14
$t-$D80M7	0.12	0.06	0.11	0.04	0.09	0.03
$t-$D81M10	0.12	0.07	0.12	0.07	0.22	0.09

See Table 4 for the DGP. 100 replications.

5.2 Bootstrap forecasts

Extended bootstrap samples are used to check the reliability of our asymptotic likelihood-based inference on the forecasts. Note that the derivations of forecast intervals neglect the estimation uncertainty in the parameters and treat them effectively as known. It is interesting to see whether this is appropriate in the current context, where we re-estimate the model in each bootstrap sample.

We present the results for US inflation in Table 6. For each estimation method, forecast method, and each horizon we present six numbers. The first two are the actual forecast and its RMSE (as reported for the UK price level in Table 2). The next four statistics are obtained from the bootstrap simulations. The column labelled 'mean' reports the mean forecast error of the 100 replications, followed by MCSD, which reports the Monte Carlo standard deviation, i.e. the standard deviation of the 100 forecast errors. The final two statistics describe the empirical distribution of the estimated RMSEs: their mean (labelled 'RMSE mean') and standard deviation (labelled 'RMSE MCSD') . For one-step ahead forecasts these are mainly determined by the empirical distribution of $\widehat{\sigma}_{\varepsilon}$. For longer multi-step forecasts they are also influenced by $\widehat{d}$.

Overall the empirical MCSDs in column 6 compare well with the RMSE that we estimated directly using the standard formulae for our actual sample (column 4). Note that the EML-estimator of the RMSE is somewhat downward biased in comparison to EML. The uncertainty in the estimated RMSE is quite small, as shown in the final column Table 6.

We see again that the choice of estimator matters much more than the choice of predictor. In this case the difference between optimal and naive forecasts is negligeable.

6 Allowing for GARCH effects

The Gaussian homoskedastic ARFIMA-model leads to convenient chi-squared inference for d, and it also leads to normal forecast intervals. It gives an adequate

Table 6: Actual forecasts, and parametric bootstrap forecast results for US inflation up to 24 months ahead

H	Date	actual forecasts		bootstrap forecast errors			
		forecast	RMSE	mean	MCSD	RMSE mean	RMSE MCSD
			EML: optimal forecasts $\widehat{z}_{T+H}$				
1	2003.05	-0.041	0.167	-0.0110	0.169	0.162	0.0049
2	2003.06	-0.021	0.175	-0.0105	0.180	0.170	0.0053
3	2003.07	0.122	0.179	-0.0133	0.173	0.174	0.0057
12	2004.04	0.154	0.187	0.0128	0.190	0.181	0.0070
24	2005.04	0.187	0.193	0.0052	0.206	0.186	0.0082
			EML: naive forecasts $\tilde{z}_{T+H}$				
1	2003.05	-0.041	0.167	-0.0103	0.169	0.162	0.0049
2	2003.06	-0.021	0.175	-0.0091	0.182	0.170	0.0053
3	2003.07	0.123	0.179	-0.0137	0.173	0.173	0.0057
12	2004.04	0.155	0.187	0.0130	0.187	0.181	0.0070
24	2005.04	0.188	0.193	0.0052	0.206	0.186	0.0083
			MPL: optimal forecasts $\widehat{z}_{T+H}$				
1	2003.05	-0.051	0.172	-0.0093	0.176	0.172	0.0052
2	2003.06	-0.016	0.180	-0.0088	0.188	0.180	0.0057
3	2003.07	0.115	0.184	-0.0122	0.182	0.183	0.0061
12	2004.04	0.151	0.192	0.0139	0.190	0.191	0.0075
24	2005.04	0.174	0.200	0.0075	0.212	0.200	0.0093
			MPL: naive forecasts $\tilde{z}_{T+H}$				
1	2003.05	-0.051	0.172	-0.0094	0.176	0.172	0.0052
2	2003.06	-0.016	0.180	-0.0088	0.188	0.180	0.0057
3	2003.07	0.116	0.184	-0.0123	0.182	0.183	0.0061
12	2004.04	0.151	0.192	0.0139	0.190	0.191	0.0075
24	2005.04	0.174	0.200	0.0074	0.212	0.200	0.0094
			NLS: naive forecasts $\tilde{z}_{T+H}$				
1	2003.05	-0.030	0.167	-0.0290	0.172	0.165	0.00049
2	2003.06	0.003	0.176	-0.0222	0.176	0.174	0.00056
3	2003.07	0.138	0.180	-0.0260	0.167	0.177	0.00061
12	2004.04	0.188	0.189	0.0106	0.190	0.186	0.00078
24	2005.04	0.206	0.192	0.0316	0.196	0.188	0.00086

See Table 4 for the DGP; 100 replications.

characterization of the second order properties of the data and it provides optimal linear forecasts. These inference and optimality properties may not be robust. The diagnostics of both the US and UK models that we estimated reveal signs of conditional heteroskedasticity. This has been ignored up to this point, and we need to consider how it impacts on the model. After all, ARCH error models were first introduced by Engle (1982) to model the UK inflation series for the 1970s.

To start, we estimate a Gaussian ARFIMA-GARCH(1,1) model using the approximate NLS-type likelihood following Ling and Li (1997) and Baillie et al. (1996). The estimation results are presented in Table 7, and should be compared to the NLS estimates in Tables 1 and 4. Table 7 only reports the ARFIMA parameters and intercept, but the estimated models include the dummies and seasonal variables as before. Standard errors are based on the second derivatives of the log-likelihood, using zero for the cross-derivative between estimated ARFIMA (with regressors) and GARCH parameters.

Table 7: ARFIMA-GARCH estimates for UK and US inflation

	UK inflation	US inflation
$\widehat{d}$	0.555 (0.087)	0.283 (0.033)
$\widehat{\phi}_{12}$	0.659 (0.18)	0.521 (0.16)
$\widehat{\theta}_{12}$	−0.427 (0.20)	−0.434 (0.18)
Constant	−0.086 (10.9)	0.237 (0.046)
$\widehat{\alpha}_0$	0.418 (0.48)	0.000086 (0.0001)
$\widehat{\alpha}_1$	0.090 (0.067)	0.052 (0.0061)
$\widehat{\beta}_1$	0.851 (0.10)	0.944 (0.016)
$\widehat{\sigma}$	2.66	0.143

The GARCH parameters are clearly significant and indicate substantial persistence in the volatility of inflation: $\widehat{\alpha}_1 + \widehat{\beta}_1 > 0.9$ for both series, and is very close to one for US inflation. For the UK, the addition of GARCH parameters has not made much difference. For the US, however, $\widehat{d}$ has fallen by nearly two standard errors. This is somewhat in contrast to the results of Ling and Li (1997), who showed that the information with regard to d and the ARMA parameters on the one hand and the GARCH parameters on the other hand is asymptotically orthogonal, and could be the consequence of IGARCH errors. There is no qualitative difference: for the US-series we still find $d < 0.5$, and for the UK we have $d > 0.5$.

The parameter σ^2 in Table 7 denotes the unconditional variance of the innovations, or 'average' one-step-ahead forecast error variance, and is comparable between the homoskedastic models and the model with GARCH innovations. It is computed as $\sigma^2 = \alpha_0/(1 - \alpha_1 - \beta_1)$.

Standard Wald tests are unlikely to work well for this sample size and model. Therefore we consider a bootstrap test on the value of d. The hypothesis of interest is $H_0 : d = 0.5$ in the ARFIMA-GARCH model for UK inflation. The parametric bootstrap test procedure is based on the parameters estimated under H_0. Using these estimates, we generate $B = 199$ bootstrap samples, on which the unrestricted model is estimated. Following Davidson and MacKinnon (1999a) we use the H_0-parameters as starting values for the iterative estimation in each bootstrap sample, but instead we iterate until convergence. Consequently, we have B t-values on d: $P(|\hat{d}_b - 0.5| < t_b) = 0.05$ for $b = 1, \ldots, B$. The relative position of the t-value from the original model within these bootstrapped t-values gives us the bootstrap p-value. For the UK we find a one-sided bootstrap p value of 28%, which is actually close to the estimated value of 26%.

6.1 Effects of neglecting GARCH errors on inference

We extend the Monte Carlo experiment of §4 by adding (unmodelled) GARCH errors:

$$\begin{aligned}&(1-0.9L^4)(1-L)^{d_0}\left[y_t - 3.5 - 15\mathrm{VAT} - 2.5(Q_2 - Q_3)\right] = (1-0.7L^4)\varepsilon_t,\\ &\varepsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t),\\ &h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1},\end{aligned} \tag{6.1}$$

with $t = 1, \ldots, 174$, and $M = 10\,000$ replications. The GARCH-parameters are set to $\alpha_0 = 0.4, \alpha_1 = 0.2, \beta_1 = 0.85$, so that the unconditional variance is 8, which equals that used in (4.1). The results are reported in Table 8, and, as expected from the block-diagonality of the information matrix, the ARFIMA estimates are not much affected by the presence of GARCH errors when compared to Table 3. The largest difference is in σ^2, which is more downward biased in Table 8.

Table 8: Monte Carlo estimates for DGP (6.1) with $d_0 = 0.45, \alpha_0 = 0.4, \alpha_1 = 0.2, \beta_1 = 0.85$, $M = 10,000$

	EML		MPL		NLS	
	bias	RMSE	bias	RMSE	bias	RMSE
$\widehat{d}$	−0.045	0.07	−0.003	0.07	−0.011	0.08
$\widehat{\phi}_4$	−0.113	0.21	−0.060	0.17	−0.107	0.24
$\widehat{\theta}_4$	0.114	0.24	0.055	0.20	0.094	0.26
Constant	−0.085	4.74	−1.187	4.16	−0.016	5.33
VAT	0.013	1.72	0.009	1.72	0.008	1.73
$Q_2 - Q_3$	−0.001	0.55	−0.001	0.55	−0.058	4.05
$\widehat{\sigma}^2$	−1.113	2.59	−1.019	2.58	−1.083	2.59

7 Conclusions

We considered several practical issues when using ARFIMA models to model and forecast inflation. We compared three methods to estimate Gaussian ARFIMA models: exact maximum likelihood, modified profile likelihood, and nonlinear least squares. We discussed computation of the exact and modified profile likelihood.

For models that are relevant for postwar monthly US core inflation and for quarterly overall UK consumer price inflation, it was shown that MPL is clearly the preferred method of inference. A Monte Carlo analysis, and a more limited parametric bootstrap analysis revealed that it is both more reliable and more efficient. Inference on the integration parameter d is especially satisfactory. However, currently available (higher order) asymptotic approximations did not lead to reliable inference on the (trending) mean of inflation.

We also compared optimal and naive forecasts. For the sample sizes at hand the difference between the two forecasting methods turned out to be of little importance. GARCH effects were found to be important empirically, but without much effect on the estimation of the ARFIMA parameters, except for $\widehat{d}$ in the US inflation model, which was estimated to be (nearly) integrated GARCH.

Although some of the asymptotic results seem to give a good guidance for finite sample inference in our model, we recommend the application of parametric bootstrap tests. Not only can bootstrap tests be expected to provide more accurate inference, the bootstrap samples also deliver additional evidence on the adequacy of parameterizations, estimators, and corresponding Wald tests for the parameters and forecasts of ARFIMA processes for inflation series.

Colophon

We wish to thank users of the Arfima package for their comments and suggestions. Financial support from the UK ESRC (grant R000237500) is gratefully acknowledged by JAD.

Correspondence to Marius Ooms (mooms@feweb.vu.nl) Department of Econometrics, Free University Amsterdam and Tinbergen Institute Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.

References

An, S. and P. Bloomfield (1993), "Cox and Reid's modification in regression models with correlated errors", *Discussion paper*, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, U.S.A.

Baillie, R. T. (1996), "Long memory processes and fractional integration in econometrics", *Journal of Econometrics 73*, 5–59.

Baillie, R. T., C.-F. Chung, and Tieslau (1996), "Analysing inflation by the fractionally integrated ARFIMA-GARCH model", *Journal of Applied Econometrics 11*, 23–40.

Barndorff-Nielsen, O. E. and D. R. Cox (1994), *Inference and Asymptotics*, London, UK: Chapman and Hall.

Beran, J. (1992), "Statistical methods for data with long-range dependence", *Statistical Science 7*, 404–416.

Beran, J. (1994), *Statistics for Long-Memory Processes*, London: Chapman and Hall.

Beran, J. (1995), "Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models", *Journal of the Royal Statistical Society 57*, 659–672.

Beveridge, S. and C. Oickle (1993), "Estimating fractionally integrated time series models", *Economics Letters 43*, 137–142.

Bos, C. S., P. H. Franses, and M. Ooms (2002), "Inflation interval forecasting by long memory regression models", *18*, 243–264.

Brockwell, P. J. and R. A. Davis (1993), *Time Series: Theory and Methods (2nd ed.)*, USA: Springer-Verlag, New-York.

Chung, C.-F. and R. T. Baillie (1993), "Small sample bias in conditional sum-of-squares estimators of fractionally integrated ARMA models", *Empirical Economics 18*, 791–806.

Cox, D. R. and N. Reid (1987), "Parameter orthogonality and approximate conditional inference (with discussion)", *Journal of the Royal Statistical Society Series B 49*, 1–39.

Dahlhaus, R. (1989), "Efficient parameter estimation for self-similar processes", *The Annals of Statistics 17*, 1749–1766.

Dahlhaus, R. (1995), "Efficient location and regression estimation for long range dependent regression models", *The Annals of Statistics 23*, 1029–1047.

Davidson, R. and J. MacKinnon (1999a), "Bootstrap testing in nonlinear models", *International Economic Review 40*, 487–508.

Davidson, R. and J. MacKinnon (1999b), "The size distortion of bootstrap tests", *Econometric Theory 15*, 361–376.

Doornik, J. A. (2001), *Object-Oriented Matrix Programming using Ox* (4th ed.), London: Timberlake Consultants Press.

Doornik, J. A. and H. Hansen (1994), "An omnibus test for univariate and multivariate normality", *Discussion paper*, Nuffield College.

Doornik, J. A. and M. Ooms (1999), "A package for estimating, forecasting and simulating ARFIMA models: Arfima package 1.0 for Ox", *Discussion paper*, Nuffield College, Oxford.

Doornik, J. A. and M. Ooms (2003), "Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models", *Computational Statistics & Data Analysis 42*, 333–348.

Engle, R. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation", *Econometrica 50*, 987–1008.

Hauser, M. A. (1999), "Maximum likelihood estimators for ARMA and ARFIMA models: A Monte Carlo study", *Journal of Statistical Planning and Inference 80*, 229–255.

Hendry, D. F. (2001), "Modelling UK inflation, 1875–1991", *Journal of Applied Econometrics 16*, 255–275.

Hendry, D. F. and J. A. Doornik (2001), *Empirical Econometric Modelling using PcGive: Volume I* (3rd ed.), London: Timberlake Consultants Press.

Hosking, J. R. M. (1981), "Fractional differencing", *Biometrika 68*, 165–176.

Kiviet, J. F. and G. D. A. Phillips (1998), "Degrees of freedom adjustment for variance estimators in dynamic regression models", *The Econometrics Journal 1*, 44–70.

Lieberman, O. (2001), "Penalised maximum likelihood estimation for fractional Gaussian processes", *Biometrika 88*, 888–894.

Ling, S. and W. K. Li (1997), "On Fractionally Integrated Autoregressive Moving-Average models with Conditional Heteroskedasticity", *Journal of the American Statistical Association 92*, 1184–1194.

Ljung, G. M. and G. E. P. Box (1978), "On a measure of lack of fit in time series models", *Biometrika 65*, 297–303.

Odaki, M. (1993), "On the invertibility of fractionally differenced ARIMA processes", *Biometrika 80*, 703–709.

Robinson, P. (1994), "Time series with strong dependence", In C. A. Sims (Ed.), *Advances in Econometrics, Sixth World Congress*, pp. 47–95. Cambridge: Cambridge University Press.

Smith Jr, A. A., F. Sowell, and S. E. Zin (1997), "Fractional integration with drift: Estimation in small samples", *Empirical Economics 22*, 103–116.

Sowell, F. (1987), "Maximum likelihood estimation of fractionally integrated time series models", *Discussion Paper* 87-07, Department of Economics, Duke University.