An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series*

Geetesh Bhardwaj and Norman R. Swanson
Rutgers University
November 2003

Abstract

This paper addresses the notion that many fractional $I(d)$ processes may fall into the "empty box" category, as discussed in Granger (1999). We begin by showing that so-called spurious long memory may arise not only in the presence of (stochastic) structural breaks and regime switches, but also when the true data generating processes (DGPs) are linear with no structural breaks, and/or regime switching properties. However, we also present ex ante forecasting evidence based on an updated version of the absolute returns series examined by Ding, Granger and Engle (1993) that suggests that ARFIMA models estimated using a variety of standard estimation procedures yield "approximations" to the true unknown underlying DGP that provide significantly better out-of-sample predictions than AR, MA, and random walk models, with no models being "better" than ARFIMA models, based on analysis of point mean square forecast errors (MSFEs) as well as using Diebold and Mariano (DM: 1995) predictive accuracy tests. We also provide ex ante prediction evidence that 130 of the 215 macroeconomic variables used in Stock and Watson (SW: 2002) have lower (or equal) point MSFEs for ARFIMA-best models than for a large class of non-ARFIMA models. This result is obtained even though our ex-ante forecasting experiments using these variables involve prediction periods of only 200-250 months, and even though there is clear and substantial evidence of severe bias in various in-sample estimates of the differencing parameter. Furthermore, when a new prediction based estimator of d that we introduce is used to construct ARFIMA models for a subset of the SW dataset, many point MSFE error improvements associated with the use of ARFIMA models also become significant. These findings, taken together with the results of an empirical analysis of stock returns for 5 different countries and an extensive Monte Carlo analysis of a variety of different estimation and testing methods lead us to conclude that ARFIMA models might not fall into the "empty box" category after all, although much further research is needed before conclusive evidence in either direction can be given.

----

1 Introduction

The last 2 decades of macroeconomic and financial economic research have resulted in a vast array of important contributions in the area of long memory modelling, both from a theoretical and empirical perspective. From a theoretical perspective, much effort has focussed on issues of testing and estimation, and a very few important contributions include Granger (1980), Granger and Joyeux (1980), Hosking (1981), Geweke and Porter-Hudak (1983), Lo (1991), Sowell (1992a,b), Ding, Granger and Engle (1993), Cheung and Diebold (1994), Robinson (1995), Engle and Smith (1999), Diebold and Inoue (2001), Breitung and Hassler (2002), and Dittman and Granger (2002). The empirical analysis of long memory models has seen equally impressive treatment, including studies by Diebold and Rudebusch (1989, 1991a,b), Hassler and Wolters (1995), Hyung and Franses (2001), Bos, Franses and Ooms (2002), Chio and Zivot (2002), and van Dijk, Franses and Paap (2002), to name but a few.[1] The impressive array of papers on the subject is perhaps not surprising, given that long memory models in economics is one of the many important areas of research that has stemmed from seminal contributions made by Clive W.J. Granger (see e.g. Granger (1980) and Granger and Joyeux (1980)). Indeed, in the write-up disseminated by the Royal Swedish Academy of Sciences upon announcement that Clive W.J. Granger and Robert F. Engle had won the 2003 Nobel Prize in Economics, it was stated that:[2]

Granger has left his mark in a number of areas. [other than in the development of the concept of cointegration] His development of a testable definition of causality (Granger (1969)) has spawned a vast literature. He has also contributed to the theory of so-called long-memory models that have become popular in the econometric literature (Granger and Joyeux (1980)). Furthermore, Granger was among the first to consider the use of spectral analysis (Granger and Hatanaka (1964)) as well as nonlinear models (Granger and Andersen (1978)) in research on economic time series.

This paper attempts to add to the wealth of literature on the topic by asking a number of new questions, and presenting some perhaps surprising new empirical evidence on the subject.

First, as pointed out by many authors, including Diebold and Inoue (2001), Engle and Smith (1999), and Granger and Hyung (1999), so-called spurious long memory may arise in many contexts, such as the presence of (stochastic) structural breaks in linear and nonlinear models, the prevalence of regime switching models, and simple nonlinear transformation of "short memory" variables. This point has been illustrated convincingly using theoretical, empirical, and experimen-

---

[1] Many other empirical and theoretical studies are referenced in the entensive survery paper by Baillie (1996).

[2] see list of references under "Bank of Sweden (2003)" for a reference to the document.

tal arguments. We add to the discussion by showing via Monte Carlo studies that spurious long memory may arise if reliance is placed on any of 5 standard tests of short memory, even if the true data generating processes (DGPs) are linear with no data transformation, structural breaks, and/or regime switching properties. Our experimental findings are supported using an empirical analysis of three large financial and macroeconomic datasets from Engle, Granger and Ding (1993), Stock and Watson (2002) and Leybourne, Harris and McCabe (2003), where it is shown, for example, that standard short memory tests find ample evidence of long memory, even when ex ante prediction analysis indicates that relatively simple ARFIMA models constructed using 4 different estimators of the differencing parameter, d, are "bested" by simple linear AR, ARIMA, and related models, based on mean square out-of-sample forecast error (MSFE) measures and based on application of the Diebold and Mariano (DM: 1995) predictive accuracy test.

Second, there has been little evidence in the literature as to the usefulness of long memory models. For example, Granger (1999) discusses outliers, breaks, and distributional properties in the context of long memory models, concluding that there is a good case to be made for $I(d)$ processes falling into the "empty box" category. We attempt to stem the tide of negative evidence by presenting ex ante forecasting evidence based on an updated version of the absolute returns series examined by Ding, Granger and Engle (1993) and Granger and Ding (1996) that suggests that ARFIMA models estimated using a variety of standard estimation procedures yield "approximations" to the true unknown underlying DGP that provide significantly better out-of-sample predictions than AR, MA, and random walk models, with no models being "better" than ARFIMA models, based on analysis of point mean square forecast errors (MSFEs) as well as using Diebold and Mariano (DM: 1995) predictive accuracy tests.[3] Furthermore, the samples used in our ex ante analysis appear to be sufficient to remedy finite sample bias properties of 4 d-estimators (including, for example, the oft used Geweke, Porter-Hudak (GPH: 1983) and Whittle estimators) that have been so widely discussed in the literature. However, the well known bias properties are starkly prevalent for samples even as great as 2500 observations, as evidenced by our analysis of stock index returns for 5 different countries. Nevertheless, even in these cases, ARFIMA models MSFE-dominate a variety of linear models based on various returns transformations and using various sub-samples of the data. Further, based on smaller samples of 5000 daily stock index observations

---

[3]One implication of our findings is that even if there are structural breaks, etc., then ARFIMA models may be more robust to them than various parsimonious linear non-ARFIMA models.

for 5 different countries, ARFIMA-best models have MSFEs that are lower (or equal) to values for non-ARFIMA-best models in 39 of 45 cases studied. Finally, our most surprising evidence may be that 130 of the 215 macroeconomic variables used in Stock and Watson (2002) have lower (or equal) point MSFEs for ARFIMA-best models. This, even though our ex-ante forecasting experiments using these variables involve prediction periods of only 200-250 months. In summary, these results suggest that if we believe all forecasting models should be viewed as "approximations" (and are hence misspecified), then ARFIMA models seem to be an "approximation" worth including in our empirical toolbox, even though some of the theoretical memory characteristics of these models are suspect, and large samples are needed for estimation, as discussed in Granger (1999).

Third, even though we have evidence based on examination of point MSFEs that ARFIMA models can dominate well known linear (and short memory) alternatives, there is little confidence interval evidence (based on the use of DM tests) that there is anything to choose between them. One reason for this is clearly that d is poorly estimated using small samples, as shown in our empirical and Monte Carlo analyses. Indeed, our empirical evidence suggests that samples as large as 10,000 observations are required in order to acheive reasonably unbiased estimators of d. In order to address this problem, we propose a simple prediction based estimator, which is essentially constructed by carrying our ex-ante prediction experiments for various values of d in a given "feasible set". We show via a small empirical analysis of a subset of the data used by Stock and Watson (2002) that the estimator can dramatically increase the incidence of ARFIMA significantly outperforming competitor non-ARFIMA models. Theoretical analysis of the estimator, however, is left to future research.

Finally, we address, using an extensive empirical and Monte Carlo analysis, a number of other issues, such as the prevalence of parameter estimation error when estimating ARFIMA models using small samples, the usefulness of prediction based analysis for testing the null hypothesis of short memory, and what impact GARCH errors have on standard in-sample as well as out-of-sample tests of short memory.

The rest of the paper is organized as follows. In Section 2 we briefly review the ARFIMA process, and outline the empirical estimation and testing methodology that we employ in the rest of the paper. In Section 3 we present the results of our empirical investigation of more than 200 macroeconomic and financial variables. Section 4 contains the results of a series of Monte Carlo experiments that were designed to yield further evidence on a number of issues and findings based

on our empirical analysis. Section 5 contains some concluding remarks.

2 Empirical Methods

The prototypical ARFIMA model examined in the literature is

$$\Phi\left(L\right)\left(1-L\right)^{d}y_{t}=\Theta\left(L\right)\epsilon_{t} \tag{1}$$

where d is the fractional differencing parameter, ϵ_{t} is white noise, and the process is covariance stationary for $-0/5 < d < 0.5$ and mean reverts when $d < 1$. This model is a generalization of the fractional white noise process described in Granger (1980), Granger and Joyeux (1980), and Hosking (1981), where, for the purpose of analyzing the properties of the process, $\Theta\left(L\right)$ is set equal to unity (Baillie (1996) surveys numerous papers that have analyzed the properties of the ARFIMA process). Given that many time series exhibit very slowly decaying autocorrelations, the potential advantage of using ARFIMA models with hyperbolic autocorrelation decay patterns when modelling economic and financial times series seems clear (as opposed to models such as ARMA processes that have exponential or geometric decay). The potential importance of the hyperbolic decay property can be easily seen by noting that

$$(1-L)^{d}=\sum_{j=0}^{\infty}(-1)^{j}\binom{d}{j}(L)^{j}=1-dL+\frac{d(d-1)}{2!}L^{2}-\frac{d\left(d-1\right)\left(d-2\right)}{3!}L^{3}+\cdots=\sum_{j=0}^{\infty}b_{j}\left(d\right), \tag{2}$$

for any $d > -1$.[4] As a simple illustration, Table 1 reports the values of the coefficients associated with different lags in the expansion of $(1-L)^{d}$ given in equation (2). The last column of the table gives the lag after which coefficients of the polynomial become smaller than 1.0e-004. It is interesting to note that by this crude yardstick the coefficients are significant even after 400 lags, in the case when $d = 0.25$. Even when $d = 0.05$, properly capturing the dependence structure requires 300 lags.

---

[4]For $d > 0$, the differencing filter can also be expanded using hypergeometric functions, as follows: $(1-L)^{d} = \Gamma(-d)\sum_{j=0}^{\infty}L^{k}\Gamma(j-d)/\Gamma(j+1) = F(-d,1,1,L)$, where $F(a,b,c,z) = \Gamma(c)/[\Gamma(a)\Gamma(b)]\sum_{j=0}^{\infty}z^{j}\Gamma(a+j)\Gamma(b+j)/[\Gamma(c+j)\Gamma(j+1)]$

4

The Long-Memory Filter $(1 - L)^{d}$ [*]

d	lag = 5	lag = 10	lag = 20	lag = 25	lag = 50	lag = 75	lag = 100	Significant Lag
0.05	-0.00899	-0.00433	-0.00209	-0.00165	-0.0008	-0.00052	-0.00039	361
0.25	-0.0282	-0.01166	-0.00486	-0.00367	-0.00154	-0.00093	-0.00065	444
0.45	-0.02894	-0.01022	-0.00368	-0.00265	-0.00096	-0.00053	-0.00035	237
0.55	-0.02528	-0.00823	-0.00275	-0.00194	-0.00066	-0.00035	-0.00022	167
0.75	-0.01428	-0.00394	-0.00113	-0.00076	-0.00022	-0.00011	-0.00007	78

[*] Notes: Values taken by the filter $(1 - L)^{d}$ are reported in columns 2 to 8. The last column gives the lag after which the absolute value of coefficients of the polynomial become smaller than $1.0e - 004$.

There are currently dozens of estimation methods for and tests of long memory models. Perhaps one of the reasons for the wide array of tools for estimation and testing is that the current consensus on estimation is somewhat negative, and many of the tests used for long memory have been shown via finite sample experiments to perform quite poorly. Much of this evidence has been reported in the context of comparing one or two classes of estimators/tests, such as rescaled range (RR) type estimators (as introduced by Hurst (1951) and modified by Lo (1991), for example) and log periodogram regression estimators due to Geweke and Porter-Hudak (1983: GPH). In the face of all of the negative publicity, it a bit surprising to us that few papers seem to compare more that one or two different (classes of) estimators and/or tests. Our approach, while still far from exhaustive, is to use a variety of the most widely used estimators and tests in our subsequent empirical investigation and experimental analysis. In particular, we consider 4 of the most widely used estimation methods and 5 different long memory tests.[5]

2.1 Long Memory Model Estimation

The GPH estimation procedure is a two-step procedure, which begins with the estimation of d, and is based on the following log-periodogram regression[6]:

$$\ln\left[I\left(\omega_j\right)\right] = \beta_0 + \beta_1 \ln\left[4\sin^2()\omega_j 2\right] + \nu_j, \tag{3}$$

where

$$\omega_j = \frac{2\pi j}{T}, \ j = 1, 2, ..., m.$$

---

[5]Perhaps the most glaring ommission from our list of estimators is the full information maximum likelihood estimator of Sowell (1992a). While his estimator is theoretically appealing, it is computationally demanding as it requires inversion of TxT matrices of nonlinear functions of hypergeometric functions. For evidence on the finite sample performance of this estimator, the reader is referred to Cheung and Diebold (1994).

[6]The regression model is usually estimated using least squares.

The estimate of d, say $\widehat{d}_{GPH}$, is $-\widehat{\beta}_1$, ω_j represents the $m = \sqrt{T}$ Fourier frequencies, and $I(\omega_j)$ denotes the sample periodogram defined as

$$I(\omega_j) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} y_t e^{-\omega_j t}\right|^2 \tag{4}$$

The critical assumption for this estimator is that the spectrum of the ARFIMA(p,d,q) process is the same as that of an ARFIMA(0,d,0) process (the spectrum of the ARFIMA(p,d,q) process in (1), under some regularity conditions, is given by $I(\omega_j) = z(\omega_j)(2\sin()\omega_j 2)^{-2d}$, where $z(\omega_j)$ is the spectrum of an ARMA process). We use $m = \sqrt{T}$, as is done in Diebold and Rudebusch (1989), although the choice of m when ϵ_t is autocorrelated can heavily impact empirical results (see Sowell (1992b) for discussion). Robinson (1995a) shows that $(\frac{\pi^2}{24m})^{-1/2}\left(\widehat{d}_{GPH} - d\right) \to N(0,1)$, for $-1/2 < d < 1/2$, for $j = l, ..., m$ in the equation for ω above, where l is analogous to the usual lag truncation parameter. As is also the case with the next two estimators, the second step of the GPH estimation procedure involves fitting an ARMA model to data, given the estimate of d. Agiakloglou, Newbold and Wohar (1992) show that the GPH estimator has substantial finite sample bias, and is inefficient when ϵ_t is a persistent AR or MA process. Many authors assume normality of the filtered data in order to use standard estimation and inference procedures in the analysis of the final ARFIMA model (see e.g. Diebold and Rudebusch (1989,1991a). Numerous variants of this estimator continue to be widely used in the empirical literature.

Another seminparametric estimator, the Whittle estimator, is also often used to estimate d. Perhaps one of the more promising of these is the local Whittle estimator proposed by Künsch (1987) and modified by Robinson (1995b). This is another periodogram based estimator, and the crucial assumption is that for fractionally integrated series, the autocorrelation (ρ) at lag l is proportional to l^{2d-1}. This implies that the spectral density which is the Fourier transform of the autocovariance γ is proportional to $(\omega_j)^{-2d}$. The local Whittle estimator of d, say $\widehat{d}_{WHI}$ is obtained by maximizing the local Whittle log likelihood at Fourier frequencies close to zero, given by

$$\Gamma(d) = -\frac{1}{2\pi m}\sum_{j=1}^{m}\frac{I(\omega_j)}{f(\omega_j;d)} - \frac{1}{2\pi m}\sum_{j=1}^{m}f(\omega_j;d), \tag{5}$$

where $f(\omega_j;d)$ is the spectral density (which is proportional to $(\omega_j)^{-2d}$). As frequencies close to zero are used, we require that $m \to \infty$ and $\frac{1}{m} + \frac{m}{T} \to 0$, as $T \to \infty$. Taqqu and Teverovsky (1997)

6

show that $\widehat{d}_{WHI}$ can be obtained by maximizing the following function:

$$\widehat{\Gamma}(d) = \ln\left(\frac{1}{m}\sum_{j=1}^{m}\frac{I(\omega_j)}{\omega_j^{-2d}}\right) - 2d\frac{1}{m}\sum_{j=1}^{m}\ln(\omega_j) \tag{6}$$

Robinson (1995b) shows that for estimates of d obtained in this way, $(4m)^{1/2}\left(\widehat{d}_{WHI} - d\right) \rightarrow N(0,1)$, for $-1/2 < d < 1/2$. Taqqu and Teverovsky (1997) study the robustness of standard, local, and aggregated Whittle estimators to non-normal innovations, and find that the local Whittle estimator performs well in finite samples. Shimotsu and Phillips (2002) develop an exact local Whittle estimator that applies throughout the stationary and nonstationary regions of d, while Andrews and Sun (2002) develop an adaptive local polynomial Whittle estimator in order to address the slow rate of convergence and associated large finite sample bias associated with the local Whittle estimator.

The rescaled range estimator was originally proposed as a test for long-term dependence in the time series. The statistics is calculated by dividing range with standard deviation:

$$\widehat{Q}_T = \frac{\widehat{R}_T}{\widehat{\sigma}_T}, \tag{7}$$

where $\widehat{\sigma}_T^2$ is the usual maximum likelihood variance estimator of y_t, and $\widehat{R}_T = \max_{0<i\leq T}\sum_{t=1}^{i}(y_t - \overline{y}) - \min_{0<i\leq T}\sum_{t=1}^{i}(y_t - \overline{y})$. The estimate of d, say $\widehat{d}_{RR}$, is obtained using the result that $plim_{T\to\infty}(T^{-d-\frac{1}{2}}\frac{\widehat{R}_T}{\widehat{\sigma}_T})$ $= constant$ (see Hurst (1951), Lo (1991), and the references cited therein), and is:

$$\widehat{d}_{RR} = \frac{\ln\left(\widehat{Q}_T\right)}{\ln(T)} - \frac{1}{2}. \tag{8}$$

Lo (1991) shows that $T^{-1/2}\widehat{Q}_T$ is asymptotically distributed as the range of a standard Brownian bridge. Given the extensively documented deficiencies of long memory tests based on $T^{-1/2}\widehat{Q}_T$, particularly in the presence of data generated by a short memory processes combined with a long memory component, Lo (1991) suggests the modified RR test, whereby $\widehat{\sigma}_T^2$ is replaced by a heteroskedasticity and autocorrelation consistent variance estimator, namely:

$$\widehat{\sigma}_T^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \overline{y})^2 + \frac{2}{T}\sum_{j=1}^{q}w_j(q)\left\{\sum_{t=j+1}^{T}(y_t - \overline{y})(y_{t-j} - \overline{y})\right\}, \tag{9}$$

where

$$w_j(q) = 1 - \frac{j}{q+1}, q < T,$$

7

It is known from Phillips (1987) that $\widehat{\sigma}_T^2$ is consistent when $q = O(T^{1/4})$, at least in the context of unit root tests, although choosing q in the current context is a major difficulty. This statistic still weakly converges to the range of a Brownian bridge. The finite sample properties of tests based on $\widehat{Q}_T$ have been examined extensively (see e.g. Cheung (1993)), and the overall consensus is that they have poor finite sample properties.

The fourth estimator that we use is the approximate maximum likelihood estimator of Beran (1995). For any ARFIMA model given by equation (1), $d = m + \delta$, where $\delta \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, and m is an integer (assumed known) denoting the number of times the series must be differenced in order to attain stationarity, say:

$$x_t = (1 - L)^m y_t \tag{10}$$

Next, a value of δ is fixed, and an ARMA model is fitted to the filtered x_t data, yielding a sequence of residuals. This is repeated over a fine grid of $d = m + \delta$, and $\widehat{d}_{AML}$ is the value which minimizes the sum squared residuals. The choice of m is critical, given that the method only yields asymptotically normal estimates of the parameters of the ARFIMA model if $\delta \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, for example.

2.2 Short Memory Tests

Four of the five tests that we use when evaluating our time series are based on the above discussion, including the GPH, RR, MRR, and WHI tests, where the MRR is the modified RR test due to Lo (1991). Notice that of these, only the GPH and WHI tests are based directly upon examination of the d estimator, while the RR and MRR tests do not involve first estimating d. The fifth test that we use is the nonparametric short memory test of Leybourne, Harris and McCabe (2003: LHM). Their test that is based on the rate of decay of the autocovariance function. In particular, the null hypothesis of the test is that the data shows short memory (i.e. that $\sum_{j=0}^{\infty} |\gamma_j| < \infty$, where γ_j is the autocovariance of y_t at lag j), and the test is based on the notion that one can distinguish between short and long memory via knowledge of the rate at which $\gamma_j \to 0$, as $j \to \infty$. The test statistics is

$$S_{k,T} = \frac{T^{1/2}\widehat{\gamma}_{k_T}}{\widehat{\sigma}_{\infty}}, \tag{11}$$

where $\widehat{\sigma}_{\infty}^2 = \widehat{\gamma}_0^2 + 2\sum_{j=1}^{l_T} \widehat{\gamma}_j^2$, $\widehat{\gamma}_j = T^{-1}\sum_{t=j+1}^{T} y_t y_{t-j}$, y_t in this case is the demeaned series, and k_T, l_T are chosen such that $k_T, l_T \longrightarrow \infty$, as $T \longrightarrow \infty$ and $\frac{k_T}{l_T} \longrightarrow 0$, $k_T < l_T$. The values which we

use, as suggested by LHM are $k_T = \frac{5.5T^{1/2}}{\ln(T)}$ and $l_T = 4\left(\frac{T}{100}\right)^{1/4}$. In this context, $S_{k_T} \longrightarrow N(0,1)$, under the null hypothesis. There are many other important tests available in the literature which are not examined here, including but not limited to the KPSS test (see Lee and Schmidt (1996)) and the augmented Dickey-Fuller test (see Diebold and Rudebusch (1991b)).

2.3   Predictive Accuracy Testing

If, as is often the case, the ultimate goal of an empirical investigation is the specification of predictive models, then a natural tool for testing for the presence of long memory is the predictive accuracy test. In this case, if an ARFIMA model can be shown to yield predictions that are superior to those from a variety of alternative linear (and nonlinear) models, then one has direct evidence of long memory, at least in the sense that the long memory model is the best available "approximation" to the true underlying DGP. Conversely, even if one finds evidence of long memory via application of the tests discussed above, then there is little use specifying long memory models if they do not outpredict simpler alternatives. There is a rich recent literature on predictive accuracy testing, most of which draws in one way or another on Granger and Newbold (1986), where simple tests comparing mean square forecast errors (MSFEs) of pairs of alternative models under assumptions of normality are outlined. Perhaps the most important of the predictive accuracy tests that have been developed over the last 20 years is the Diebold and Mariano (1995: DM) test. It is this test statistic that we consider in the remainder of this paper. The relevant statistic is:

$$\widehat{d}_P = P^{-1/2}\frac{\sum_{t=R-h+1}^{T-1}(f(\widehat{v}_{0,t+h}) - f(\widehat{v}_{k,t+h}))}{\widehat{\sigma}_P}, \tag{12}$$

where R denotes the estimation period, P is the prediction period, f is some generic loss function, $h \geq 1$ is the forecast horizon, $\widehat{v}_{0,t+h}$ and $\widehat{v}_{k,t+h}$ are h-step ahead prediction errors constructed using consistent estimators, and $\widehat{\sigma}_P^2$ is defined as

$$\widehat{\sigma}_P^2 = \frac{1}{P}\sum_{t=R-h+1}^{T-1}(f(\widehat{v}_{0,t+h}) - f(\widehat{v}_{k,t+h}))^2 + \frac{2}{P}\sum_{j=1}^{l_P}w_j\sum_{t=R-h+1+j}^{T-1}(f(\widehat{v}_{0,t+h}) - f(\widehat{v}_{k,t+h}))(f(\widehat{v}_{0,t+h-j}) - f(\widehat{v}_{k,t+h-j})) \tag{13}$$

where $w_j = 1 - \frac{j}{l_P+1}$, $l_P = o(P^{1/4})$. The hypotheses of interest are

$$H_0 : E(f(v_{0,t+h}) - f(v_{k,t+h})) = 0,$$

and

$$H_A : E(f(v_{0,t+h}) - f(v_{k,t+h})) \neq 0.$$

9

The DM test, when constructed as outlined above, has a standard normal limiting distribution under the null hypothesis. West (1996) shows that when the out-of-sample period grows at a rate not slower than the rate at which the estimation period grows (i.e. $\frac{P}{R} \to \pi$, with $0 < \pi \le \infty$), parameter estimation error generally affects the limiting distribution of the DM test in stationary contexts. On the other hand, if $\pi = 0$, then parameter estimation error has no effect. Additionally, Clark and McCracken (2001) point out the importance of addressing the issue of nestedness when applying DM and related tests.[7] Other recent papers in this area include Christoffersen (1998), Christoffersen and Diebold, (1997), Clements and Smith, (2000,2002), Corradi and Swanson, (2002), Diebold, Gunther and Tay (1998), Diebold, Hahn and Tay, (1999), Harvey, Leybourne and Newbold (1998), and the references contained therein, to name but a few. Although the DM test does not have a normal limiting distribution under the null of non causality when nested models are compared, the statistic can still be used as an important diagnostic in predictive accuracy analyses. Furthermore, the nonstandard limit distribution is reasonably approximated by a standard normal in many contexts (see McCracken (1999) for tabulated critical values). For this reason, we use standard normal critical values when assessing our DM statistics, even though some of our models are nested.

2.4 Prediction Based Estimator of d

As discussed in Granger (1999), the distinguishing feature of long memory processes is at very low spectral frequencies, and information accumulates very slowly there. Indeed, the empirical and experimental evidence presented below suggests that the different estimators of d that we employ in our analysis can be vastly different, only appearing to converge to one another when huge time series of around 10,000 observations are used for our in-sample estimation. This small sample bias is problematic from an empirical perspective, and may account in part for the fairly widespread dissatisfaction with ARFIMA processes in the context of modelling macroeconomic variables.

If the objective of an empirical investigation is the construction of the "best" prediction models, then a natural and simple estimator of d would be to split the sample into in- and out-of-sample

---

[7]Chao, Corradi, and Swanson (2001) address not only nestedness, by using a consistent specification testing approach to predictive accuracy testing, but also allow for misspecification amongst competing models; an important feature if one is to presume that all models are approximations, and hence all models may be (dynamically) misspecified. White (2000) further extends the Diebold and Mariano framework by allowing for the joint comparison of multiple models, while Corradi and Swanson (2003) extend White (2000) to predictive density evaluation with parameter estimation error.

period. Then, for a given value of d, estimate the model and construct a sequence of ex-ante h-step ahead predictions. Repeat this across a fine grid of different d values, and use the value of d that minimizes MSFE as the prediction-best estimator of d. While the asymptotic equivalence of this estimator with the other estimators above may be difficult to show for values of d close to or greater than 0.5, say, the estimator should nevertheless be useful as it may reduce finite sample bias. Additionally, the estimator is robust to the lag truncation procedure used in constructing $(1 - L)^d$. For a given lag truncation procedure, the estimator simply yields the MSFE best value of d. Analysis of the asymptotic properties of this estimator are the subject of ongoing research. Evidently, however, the most substantial gains to using the prediction based estimator may be when time series are relatively short, such as is the case with most macroeconomic time series. For this reason, our experimentation using this estimator involves analysis of a subset of the Stock and Watson (2002) dataset (see Section 3.3).

In the sequel, all predictions are ex-ante 1-step ahead forecasts, models are estimated once using R in-sample observations, $R = 0.5T$, so that $1/2$ of the sample observations are held back for ex ante predictive analysis, and reported DM tests are based on MSFE loss. Results for mean absolute deviation and mean absolute percentage error loss functions have also been tabulated, and are available upon request from the authors.

3 Empirical Evidence

In our empirical (and subsequent Monte Carlo) investigation, the following models are used:

1) *ARFIMA(p,d,q)*: $\Phi(L)(1 - L)^d y_t = \alpha + \Theta(L)\epsilon_t$, where d can take fractional values;

2) *Random Walk*: $y_t = y_{t-1} + \epsilon_t$;

3) *Random Walk with Drift*: $y_t = \alpha + y_{t-1} + \epsilon_t$;

4) *AR(p)*: $\Phi(L)y_t = \alpha + \epsilon_t$;

5) *MA(q)*: $y_t = \alpha + \Theta(L)\epsilon_t$;

6) *ARMA(p,q)*: $\Phi(L)y_t = \alpha + \Theta(L)\epsilon_t$;

7) *ARIMA(p,d,q)*: $\Phi(L)(1 - L)^d y_t = \alpha + \Theta(L)\epsilon_t$, where d can take integer values;

8) *ARX(p)*: $\Phi(L)y_t = \alpha + \beta X_t + \epsilon_t$;

9) *GARCH*: $\Phi(L)y_t = \alpha + \epsilon_t$, where $\epsilon_t = h_t^{1/2}\nu_t$ with $E(\epsilon_t^2|\Im_{t-1}) = h_t = \varpi + \alpha_1\epsilon_{t-1}^2 + \cdots + \alpha_q\epsilon_{t-q}^2 + \beta_1 h_{t-1} + \cdots + \beta_p h_{t-p}$, and where $\Im_{t-1}$ is the usual filtration of the data; and

10) *Regime Switching:* $y_t = \mu_{s_t} + \epsilon_t$,

where $\{s_t\}_{t=1}^T$ is the state vector with transition matrix $P = \begin{pmatrix} p_{00} & 1 - p_{00} \\ p_{11} & 1 - p_{11} \end{pmatrix}$. In these models, ϵ_t is the disturbance term, $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$, and $\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q$, where L is the lag operator. All models (except ARFIMA models) are estimated using (quasi) maximum likelihood, with values of p and q chosen via use of the Schwarz Information Criterion (SIC), and integer values of d in ARIMA models selected via application of the augmented Dickey Fuller test (using a 5% nominal test level). Errors in the GARCH models are assumed to be normally distributed. ARFIMA models are estimated using the four estimation techniques discussed above (GPH, RR, WHI, and AML). In this section, we omit the regime switching (as the model is too simplistic) models, although these models are considered in selected Monte Carlo experiments. Additionally, note that the ARX models are meant as a strawman model, as X_t is an $iid N(0, 1)$ random variable. When fitting ARFIMA models, we used an arbitrary cut-off of $1.0e - 004$. Terms in the polynomial expansion with coefficients smaller in absolute value than this cut-off were truncated.[8]

In the proceeding sub-sections, we carry out our empirical investigation by examining the long memory and ARFIMA predictive properties of the S&P500 series used by Ding, Granger and Engle (DGE: 1993) and Granger and Ding (1996), the 5 stock index returns used by Leybourne et al. (2003), and the 215 Stock and Watson (2002) macroeconomic variables. Before discussing the results, however, some comments concerning the data are in order. Our first dataset is an updated version of the long historical S&P500 returns dataset of DGE. The period covered is January 4, 1928 - September 30, 2003 (20,105 observations), so that our dataset is somewhat longer than the 17,054 observations (ending on August 30, 1990) examined by DGE. Our second dataset consists of the returns data used in Leybourne et al. (2003), where strong evidence of long memory is found via application of their short memory test. In particular, we model 4950 (or more, depending on the particular index) daily returns for the following stock indexes: S&P500, FTSE100, DAX, Nikkei225, and the Hang Seng. We consider absolute returns, squared returns, and log squared returns, thus nesting a variety of different data transformations that have been shown in earlier papers (see e.g. Granger and Ding (1996)) to have long memory properties. All series span the period 01/04/1981-01/18/2002. Finally, we examine the Stock-Watson dataset, which consists of

---

[8]The exception to the rule is the case of the SW data, for which sufficient observations were not available, and for which, after some experimentation, the arbitrary cut-off was set at 120 lags.

the 215 variables used in their well known diffusion index paper. In the paper, they examine multi-step ahead predictions of 8 key U.S. macroeconomic variables, in a simulated real-time forecasting environment, using all 215 U.S. series to construct diffusion indexes. Their data were collected in 1999, and so represent a snapshot of the vintages and releases of data available at that point in time. The data series vary in length, span the period 1959-1998, and are generally 400-500 observations in length. All series are monthly. Appendix 2 of Stock and Watson (2002) contains definitions of all of the series (which are omitted here for the sake of brevity), and discusses the data transformations applied to each series. Note that all series were "differenced to stationarity" in Stock and Watson (2002), prior to model fitting. We use the same data transformations as they did, so that many of the series are expressed in growth rates, and some series are even differenced twice. In summary, our approach is to use exactly the same dataset as used in Stock and Watson (2002). However, rather than focusing on predictions of 8 series, we consider predictions of all 215 variables, using estimated versions of the models outlined above.

3.1 S&P500 Returns: January 4, 1928 - September 30, 2003

The table given below contains d estimates, as well as MSFEs for all of the models given above, with the exception of the RW with drift and ARIMA models. These two models are not considered because the first model is nonsensical and the second model is not appropriate given the ample evidence in the literature that $d<0.5$ for absolute S&P500 returns. Interestingly, and as might be expected, the 4 different estimates of d given in the table are all quite close to one another, perhaps with the exception of the RR estimator, which is 0.31. In the next subsection, we show that these 4 estimators yield radically different values with even when the in-sample period used is moderately large, with approximately 2500 observations, so that the convergence of the estimators is indeed extremely slow, although they do eventually converge. This yields credence to Granger's (1999) observation that estimates of d can vary greatly across different sample periods and sample sizes, and are generally not robust at all (see next section for further evidence of this).[9]

Turning next to the MSFEs reported in the table, a number of clear-cut conclusions emerge. First, notice in the third column of entries that the MSFEs associated with predictions from the 4 different ARFIMA models are all very close to one another, so that based on point MSFEs, there

---

[9]Notice that the estimates of d in the table that are based on the GPH estimator are very close to those obtained by Ding, Granger and Engle (1993) and by Granger and Ding (1996) using their fractionally integrated ARCH model.

appears little to choose between GPH, AML, RR, and WHI. However, the DM test statistic does prefer the RR estimator over the other 3, at a 20% level (this is not surprising, given the tiny standard errors associated with construction of the DM statistic using such a large out-of-sample period - recall that approximately 10,000 forecasts were made for each model). Second, there is no instance for which one of our linear models out-predicts any of our ARFIMA models, both based on point MSFE values and based on DM predictive accuracy tests (except for ARMA, which leaves nothing to choose between the two models). Furthermore, the ARFIMA model estimated using RR significantly out-predicts every linear model, with the exception of the ARMA model, for which there is nothing to choose between the two. Finally, the ARMA model significantly outpredicts all of the other linear alternatives. These results constitute evidence that long memory models may actually be useful, particularly when constructing forecasting models. Furthermore, as long as the in-sample period is very large, then all of our differencing operator estimators perform adequately, and any one of them can be successfully used to estimate "winning" prediction models. Put differently, no model from amongst those considered performs better than our simple ARFIMA models. It should, however, be stressed that structural breaks, regime switching, etc. have not been accounted for in any of our models, and it remains to see whether the types of results obtained here will also hold when structural breaks and regime switching are allowed for in both our short memory and long memory models. Some results in this regard are given in the next subsection, where different return series are examined both pre- and post-1987.

S&P500 Daily Absolute Returns - Prediction Using Different ARFIMA Estimators[*]

Model Variety	d Values and Lags Chosen	MSFE	Comparison to RR Using DM Test	Comparison to ARMA Using DM Test
GPH	ARFIMA(1,0.49,1)	0.539	-1.62, RR	-0.44, NONE
AML	ARFIMA(1,0.41,1)	0.538	-1.29, RR	-0.78, NONE
RR	ARFIMA(1,0.31,3)	0.537	——	0.37, NONE
WHI	ARFIMA(1,0.43,1)	0.538	-1.31, RR	-0.13, NONE
RW	——	1.000	-13.3, RR	-11.1, ARMA
AR	AR(11)	0.549	-4.15, RR	-3.75, ARMA
MA	MA(35)	0.555	-5.49, RR	-5.93, ARMA
ARX	ARX(12	0.550	-4.40, RR	-4.36, ARMA
ARMA	ARMA(4,4)	0.538	-0.37, NONE	——
AR-GARCH	AR(6)-GARCH(1,1)	0.629	-6.22, RR	-5.86, ARMA

[*] Notes: Models are estimated as discussed above, and model acronyms used are as outlined in Table 1. MFSEs are normalized by the RW MSFE. All predictions are 1-step ahead, the out-of-sample period is equal to 50% of the sample, and Diebold and Mariano (DM) tests are based on MSFE loss, and application of the test assumes that parameter estimation error vanishes and that the standard normal limiting distribution is asymptotically valid, as discussed in Section 2.3. Tests are based on nominal 20% level critical values. "Winning" models are given beside DM statistic values, when there is a "winner".

3.2 International Stock Index Returns: January 4, 1981 - January 18, 2002

Table 2 contains our results based on empirical analysis of 5 different stock market indices. Absolute, squared, and log squared returns are tested using the 5 short memory tests discussed above, ARFIMA and non-ARFIMA models are estimated, and ex-ante predictions are made and compared. As discussed above, the out-of-sample period consists of 1/2 of the sample. A number of conclusions can be made based on the analysis reported in the table. First, notice in column 4 that the short memory null hypothesis is rejected most of the time, for most of the indexes, regardless of how returns are transformed prior to test statistic construction. At face value, this might be taken as strong evidence of the potential usefulness of ARFIMA models for these data. However, we shall later see (in the Monte Carlo experiments reported below), that the 5 tests used in our study have poor size properties when faced with simple linear models, such as AR processes, with reasonably large roots (such as that associated with a slope parameter of 0.75 in an AR(1) process). Thus, the tests are probably unreliable for the types of data usually examined by macroeconomic and financial economists. This is one of the reasons why we additionally include the results of our out-of-sample forecast evaluation. The entries in the 5th and 6th columns of the table include the MSFE associated with the "best" non-ARFIMA and ARFIMA models. Starred entries in the 6th column correspond to best ARFIMA point MSFE values that are lower than or equal to the non-ARFIMA counterparts reported in the 5th column. Thus, ARFIMA-best models have MSFEs that are lower (or equal) to values for non-ARFIMA-best models in 39 of 45 cases. Although this result is surprisingly positive with respect to the usefulness of ARFIMA models, notice from the last column in the table that there is actually nothing to choose between ARFIMA and non-ARFIMA models in 33 of 45 cases, based on application of the DM test at a 20% level. Thus, while the ARFIMA model are not usually "bested" by non-ARFIMA counterparts in the sense of point MSFE, they don't usually perform significantly differently. In this sense, our predictive accuracy type test finds little evidence in favor of ARFIMA models, contrary to the results of the 5 short memory tests. Nevertheless, our findings from the longer S&P500 sample suggest that many of the ARFIMA point MSFE "wins" would become significant, were we to increase sample lengths. One reason the ARFIMA forecasting models are no longer significantly better is that there are simply too few data to get reasonable estimates of d (note that huge variation in the different estimates of d reported in the 3rd column of the table). It thus appears, as expected, that many observations are

15

needed if the objective is to estimate ARFIMA models that are useful for forecasting. Alternatively, one could consider using the prediction based estimator of d. This is done below. First, however, it is worth noting that our estimated values for d seem to be much more robust to changing the data sample (compare the entire sample results with pre- and post-1987 results) than to changing the estimator used. The robustness problem often discussed in the context of long memory models may thus be as much due to the slow and differing convergence rates of the different estimators as to the presence of structural breaks and regime switching, for example. However, the importance of structural breaks and regime switching behavior cannot be overlooked, as discussed in Granger and Hyung (1999) and Diebold and Inoue (1999).

3.3 Stock-Watson Macroeconomic Dataset: 1959-1998

Table 3 contains results of our ARFIMA modelling analysis for the 215 time series examined by Stock and Watson (SW: 2002). The table is organized in the same way as Table 2. As mentioned above, the series have variously been differenced, log differenced, etc., according to the definitions given in Appendix 2 of SW. Perhaps the most important feature of the dataset is that it contains macroeconomic variables, with sample periods ranging from 1959-1998, so that only 400-500 monthly observations are available. Thus, we are subjecting the ARFIMA models to a very stringent test when using them to construct prediction models. Given that we know the estimates of d will be suspect, it would be very surprising if any ARFIMA models were shown to out-predict more parsimonious and precisely estimated AR, ARMA, and related models.

Turning to the results reported in the table, notice first that the incidence of rejection of the null hypothesis of short memory is less than was the case using our financial datasets. However, the variability of d across estimators for a given variable does not appear to increase relative to the cases where we examined much larger financial datasets. Furthermore, 130 of 215 variables have lower (or equal) point MSFEs for the ARFIMA-best models, versus 85 "wins" for the non-ARFIMA best models. This result is rather surprising, as it suggests that even though the ARFIMA models are estimated imprecisely, selecting across our 4 different d estimators is enough to result in ARFIMA models beating linear AR, ARX, ARMA, ARIMA, RW and RW with drift models, where in the linear models the parsimonious SIC is always used to select lags. Thus, even based on the very short series that we have examined, there is surprising evidence in favor of at least adding ARFIMA models to the "toolbox" of models to be tried when fitting macroeconomic variables. This said, it

should be noted that based on application of the DM test, non-ARFIMA models are significantly better (or worse) than ARFIMA models in only 63 of 215 cases (see column 6 of the table), with approximately 1/2 of the "wins" going to ARFIMA models. Thus, the caveats discussed in the previous subsection with regard to findings of significant differences between models also apply for our macroeconomic variables.

The table reported below attempts to address the very apparent severe bias of our d estimates by constructing the prediction based estimator discussed above. In particular, we again use 1/2 the sample to construct ex ante predictions, but do so for each value of d between 0 and 1, by increments of 0.01. This experiment was done only for the IP series from Table 3, all of which appear to be characterized by values of d between 0 and 1. MSFE-best d values are reported in the second column of the table, and ARFIMA models thus estimated are compared to the non-ARFIMA-best models via application of the DM test, as in our prior analyses.[10] Now, 9 of 22 variables have significantly superior ARFIMA models, while only 3 variables yield forecasts that are significantly better when non-ARFIMA models are used. This compares to only 2 significant wins for ARFIMA versus 4 for non-ARFIMA models based on our previous analysis using in-sample estimators of d (see Table 3). This result suggests that ARFIMA models can be estimated that in fact outperform simpler linear models, even when samples are very short. As discussed above, one reason for this may be that the prediction based estimator automatically adapts itself to the particular "rule" that is being used to carry out lag truncation when expanding the lag polynomial associated with the ARFIMA model. It is also interesting to note that the prediction based estimator results in values of d that are very close to zero for 7 of the variables, including IPE, IPI, IPM, IPX, IPXDCA, IPXMCA, and IPXNCA, yielding clear evidence that no fractional differencing is necessary. In contrast, estimates reported in Table 3 for these variables, while all greater than 0, range almost uniformly between 0 and 1, with very few values being particularly near to zero. We take this as more evidence of the potential usefulness of the prediction based estimator.

---

[10]It should be noted that the DM tests are thus not entirely ex ante in this case, as the ARFIMA models actually use MSFE-best d values which are estimated via the same forecasting excercise used to compare the models. To carry out a truly fair comparison, the samples in our analysis would need to be further split into two more in- and out-of-sample periods to be used in estimation, prior to ex ante forecast construction.

17

Prediction Based Estimation of $d^{(*)}$

Macroeconomic Variable	d Estimates	MSFE, best non-ARFIMA	MSFE,ARFIMA	DM Test
IP	0.211	0.66,AR(1)	0.65*	-0.83, NONE
IPC	0.961	0.46,AR(1)	0.44*	-2.19, ARFIMA
IPCD	0.971	0.49,AR(1)	0.49*	0.06, NONE
IPCN	0.921	0.44,MA(1)	0.40*	-2.16, ARFIMA
IPD	0.481	0.64,AR(1)	0.62*	-1.63, ARFIMA
IPE	0.001	0.49,ARIMA(1,1,1)	0.49*	0.63, NONE
IPF	0.301	0.53,MA(1)	0.51*	-1.36, ARFIMA
IPI	0.051	0.51,MA(3)	0.51*	0.01, NONE
IPM	0.021	0.67,AR(1)	0.67*	-0.13, NONE
IPMD	0.191	0.69,AR(1)	0.76	2.60, AR(1)
IPMFG	0.231	0.65,AR(1)	0.64*	-0.67, NONE
IPMIN	0.991	0.49,AR(1)	0.48*	-1.44, ARFIMA
IPMND	0.471	0.56,ARX	0.57	0.48, NONE
IPN	0.341	0.54,AR(2)	0.52*	-1.55, ARFIMA
IPP	0.931	0.57,MA(1)	0.54*	-1.62, ARFIMA
IPUT	0.521	0.41,MA(1)	0.38*	-2.15, ARFIMA
IPX	0.001	0.97,AR(1)GARCH	1.05	0.79, NONE
IPXDCA	0.001	0.97,AR(1)GARCH	1.12	1.63, AR(1)GARCH
IPXMCA	0.001	0.97,AR(2)	0.97*	-1.12, NONE
IPXMIN	0.871	1.00,RW-D	1.03	0.93, NONE
IPXNCA	0.001	0.98,AR(1)	1.06	1.59, AR(1)
IPXUT	0.991	0.96,AR(1)	0.91*	-1.52, ARFIMA

$^{(*)}$ Notes: Data used in this table correspond to those used in Stock and Watson (2002), and the variables are monthly, spanning the period 1959-1998. The second column gives d, based on minimum MSFE and a grid of 100 possible values from [0.001,0.991]. Starred entries correspond to cases for which the point MSFE of the ARFIMA model was lower (or the same) as for the non-ARFIMA-best model. MSFE, best non-ARFIMA reports the lowest non-ARFIMA MSFE and associated model. All MSFEs are normalized by that of the Random Walk forecasting model. The fourth column lists the normalized MSFE of the ARFIMA model using the predction based estimate of d. The last column reports DM statistics for comparing the ARFIMA model with lowest MSFE non-ARFIMA model. "NONE" is reported when neither of the two models is preferred, based on application of the test at a nominal 20% level.

4   Experimental Evidence

Table 1 summarizes the DGPs and parameterizations considered in our experiments. The DGPs are the same as those used in our empirical analysis. For the ARFIMA models, data are generated using fractional values of d over the interval $(0, 1)$, including $d = \{0.05, 0.25, 0.45, 0.55, 0.75\}$. Additionally, MA(1) and AR(1) coefficients were specified, including $\{-0.5, 0.0, 0.5\}$ (MA) and $\{0.60, 0.75, 0.95\}$ (AR). Thus, we examine forty-five different ARFIMA specifications. When generating ARFIMA data, we used an arbitrary cut-off of $1.0e-004$. Terms in the polynomial expansion with coefficients smaller in absolute value than this cut-off were truncated. All DGPs include at most one lag, so that AR models have one autoregressive lag, MA models have one moving average lag, and ARX models have one autoregressive lag and one lag of the "exogenous" variable, etc. All variables are generated using standard normal errors, and exogenous variables are also $i.i.d.$ $N(0, 1)$. Autoregressive slope parameters considered include $\{0.25, 0.50, 0.75, 0.95\}$, the coefficient on exogenous variables in ARX models is 0.5, the MA models have coefficients equal

to $\{-0.75, -0.20, 0.25, 0.75\}$, and values of d equal to 0 and 1 are considered. For the GARCH DGP, three different specifications were considered, corresponding to three different AR coefficients, $\{0.0, 0.5, 0.9\}$. In the three regime switching DGPs examined, two states $\{-1,1\}$ and transition probabilities of $\{0.950, 0.990, 0.995\}$ were specified. In all DGPs, intercept parameters are set equal to zero. Samples of $T = \{200, 500, 2000\}$ were generated, although we report results only for the latter two sample sizes, for the sake of brevity. Complete results are available upon request from the authors. All experiments are based on 500 Monte Carlo replications.

4.1 Long Memory

Tables 4 and 5 contain finite sample power and level results for the 5 short memory tests discussed above. With regard to power, it is interesting to note that even for samples of 500 observations, all tests have reasonable power against a preponderance of the ARFIMA models. The exceptions appear to be the: (i) the LHM test, for which power is extremely low (below 0.10 in many cases) when $T = 500$, $d = 0.05$ or $d = 0.25$ and when $T = 2000$, $d = 0.05$ (except when the autoregressive coefficient is 0.95); and (ii) the WHI test, for which power is very low (below 0.30 in most cases) when $T = 500$ or $T = 2000$ and $d = 0.05$. As is well known, the problem is not with power, but with size. Notice that in Table 4 rejection frequencies are often far above the nominal 5% levels used in the tests, even though none of the DGPs contain structural breaks. For example, even when $T = 2000$, the RR, MRR, GPH, and WHI tests all have rejection frequencies of around 15% or higher, when the model is an AR(1) with slope coefficient equal to 0.75. When the slope coefficient increases to 0.95, the rejection frequencies for these tests all jump to at least 0.85. Thus, it appears that spurious long memory can even arise in simple contexts where there are no structural breaks, for example. The same sorts of findings hold for our MA, GARCH, ARX, and ARIMA models. Whenever the series is even modestly persistent, various short memory tests we use are somewhat poorly sized. Of course, and as expected given the recent literature on the subject, all short memory tests have great power against regime switching models! These results serve to reaffirm earlier findings concerning the apparent non-robustness of short memory tests, although with the caveat that problems occur even when there are no structural breaks and/or regime switching features in the DGPs.

As discussed above, it is well known that long time series are necessary in order to accurately estimate d. Table 6 reports averages and standard deviations of estimated d values for our ARFIMA

DGPs across the 500 different Monte Carlo iterations, and it is clear that even samples of 2000 observations may not be sufficient to yield reasonable estimates. Interestingly, of the 4 estimators considered, the AML estimator is clearly superior, at least in terms of bias. The RR estimator has severe negative bias for values of $d¿0.25$, while the GPH and WHI estimators perform similarly, consistently overestimating the true value of d. However, it is also worth noting that the least bias estimator also has substantially higher standard errors, as might be expected. Finally, the presence of moderate MA components (see Table 1 for details of the MA components in the models) does not appear to impact estimation, while an AR coefficients above 0.75 impact a substantial positive bias on estimates, regardless of the estimator used. Thus, if one argues that virtually all economic and financial time series of interest are characterized by a high level of persistence, even when they are short memory, then the results of Table 6 serve as a warning that none of the estimators examined in this paper yield reasonable estimates of d. In summary, even if we have relatively long samples (of 2000 observations), standard estimators be have substantial bias.

Table 7 considers a different issue. Given the prevalence of spurious long memory, even if the true DGP is a simple AR or ARIMA model without breaks, say, what sorts of values for d might one expect to find in the data. The relevant DGPs to examine here are those for which the empirical level of the short memory tests discussed above is substantially above the nominal level. This includes models with AR slope coefficients of 0.75 and 0.95, for example. In such cases, as might be expected, positive values of d that are more than 2 standard deviations away from 0 are not uncommon. Indeed, mean values from around 0.2 to 0.7 are quite the norm in such cases. This supports the notion that the range of values for d often found for economic and financial time series are not only consistent with ARFIMA models with such difference parameter values, but also with simple AR and ARX models with relatively high levels of persistence, against which short memory tests appear to reject much too frequently. Interestingly, the presence of MA components of the types we consider seems to dampen estimates of d, often even yielding negative values, and usually yielding simple two standard deviation confidence intervals that cover the origin.

Given the prevalence of spurious long memory, it is of interest to ascertain when predictive ability tests may serve as a further yardstick for measuring the usefulness of long memory processes. Put another way, are estimated long memory models useful for prediction when the true DGP is long memory, and if so, will predictive accuracy tests yield a more useful measure of long memory in the cases where in-sample long memory tests perform poorly? This is a question that was answered

to some degree in the above empirical section, and is further adressed in the following subsection.

4.2 Prediction

Tables 8 and 9 report the proportion of times that various models "win" a forecasting competition, based on direct comparison of point MSFE (using sequences of 1-step ahead predictions - see above for further details). Bracketed values contain analogous entries, except that DM test rejection is required for a model to "win", rather than the simpler requirement that the point MSFE of the "winning" model is lower (test inference is carried out at a nominal 20% level).

Table 8 contains results for the case where the true DGPs are ARFIMA processes. Notice that for $T =$500, MSFE-best non-ARFIMA models (from amongst the models outlined at the top of this section) outpredict the MSFE-best ARFIMA models 30-70% of the time, based on point MSFE comparison, with the best ARFIMA performance in the range where $d = 0.25 - 0.55$. However, ARFIMA models perform somewhat better in terms of point MSFE when $T = 2000$. This occurs for DGPs with AR coefficients of 0.60 and 0.75, with $d = 0.25 - 0.55$. Incidentally, these are also the DGPs for which d is estimated much more precisely when $T = 2000$ (see Table 6). Nevertheless, for other d values and/or an AR coefficient of 0.95, even a sample size of 2000 does not yield substantial improvement in terms of point MSFE. These findings support our conclusions based on our empirical analysis. Namely, samples of even $T = 2000$ are not generally sufficient to yield reasonably unbiased estimates of d. Not surprisingly, there is often no significant difference between the MSFE-best ARFIMA and non-ARFIMA models. This is apparent when one inspects the bracketed entries in the table. The only time that the "best" ARFIMA model appears to appreciably outperform the "best" non-ARFIMA model is when $d = 0.05$ (so that virtually no differencing is necessary). In all other instances, there is little to choose between the models. This is true even when $T = 2000$. Hence, point MSFE improvements associated with $T = 2000$ are not sufficient to yield statistically significant performance improvement; more observations are needed. Naturally, the main reason for this finding is undoubtedly that parameter estimation error (PEE) plays a crucial role in long memory models. Any potential gain from correctly specifying the model is offset by PEE, and more parsimonious non-ARFIMA models serve as better approximations to the true ARFIMA process than do poorly estimated ARFIMA models. The importance of PEE is illustrated in the following table.

<div align="center">The Impact of Estimating $d^{(*)}$</div>

		Sample Size 500				Sample Size 2000			
Model	d	GPH	WHI	RR	AML	GPH	WHI	RR	AML
ARFIMA1	0.05	82.8	83.6	83.4	86.6	86.2	86.0	93.0	83.4
ARFIMA9		100.0	98.6	98.4	97.4	100.0	100.0	100.0	89.4
ARFIMA10	0.25	89.0	88.8	65.4	90.0	86.2	86.4	70.8	86.6
ARFIMA18		100.0	99.2	73.4	99.4	100.0	100.0	82.2	99.6
ARFIMA19	0.45	90.6	91.0	93.2	89.0	89.0	89.2	99.6	91.0
ARFIMA27		99.8	100.0	84.0	90.4	100.0	100.0	97.8	95.6
ARFIMA28	0.55	92.4	90.2	99.0	82.0	91.6	91.2	100.0	85.6
ARFIMA36		100.0	99.8	98.6	80.2	100.0	100.0	100.0	84.6
ARFIMA37	0.75	94.0	94.8	100.0	91.0	91.0	90.4	100.0	84.0
ARFIMA45		99.8	99.4	100.0	77.0	100.0	100.0	100.0	79.0

$^{(*)}$ Notes: MSFE based comparison between ARFIMA models with all parameters known and models where all parameters except d are known. Entries denote the percentage of times for which the MSFE (based on a sequence of 1-step ahead predictions) is lower when all parameters are known. Results are based on 500 Monte Carlo iterations.

Entries in this table for many of our ARFIMA DGPs are omitted as they are qualitatively the same. The extremely high percentages recorded in the above table are as expected, as they measure the percentage of times that models with all parameters known outperform MSFE-best models with d estimated. What is perhaps surprising is that with respect to MSFE, the impact of estimating d remains essentially unchanged, even when the sample size is increased to 2000 observations, again affirming that very long samples are needed before the impact of PEE begins to diminish.

Table 9 contains results for the case where the true DGPs are non-ARFIMA processes. In almost all cases, the point MSFE-best non-ARFIMA models dominate the point MSFE-best ARFIMA models more than 50% (and sometimes up to 90%) of the time. It is interesting to note that in the rare cases where ARFIMA models are point MSFE-best around 40% to 50% of time, the average estimated d values are in the neighborhood of zero (see Table 7), and the true DGPs are AR, ARX and ARMA DGPs with small AR coefficients. Furthermore, ARFIMA models are almost never significantly better than the "best" non-ARFIMA models (see bracketed values in the table), while non-ARFIMA models are significantly better up to around 75% of the time, and are almost always MSFE-best.[11] Interestingly, the best cases for non-ARFIMA DGPs correspond to instances where the true DGP is GARCH. In particular, when the true DGP is GARCH, then the non-ARFIMA model is significantly better up to 71% of the time when $T = 500$, and 78% of the time when $T = 2000$. This suggests that when the true model is GARCH, the DM test serves as a good tool for comparing GARCH against ARFIMA models, particularly when the AR slope parameter is non-zero, as is the case in the GARCH2 and GARCH3 models. This is an interesting

---

[11]Results for regime switching models are not reported in Table 9 due to estimation difficulties in a number of instances where the true DGP was not regime switching.

finding, given the overwhelming importance of GARCH models in empirical macroeconomics and finance. On the flip side, our results suggest that even if the true model is ARFIMA, then current estimation techniques are not adequate to warrant the use of ARFIMA models for prediction, even for $T = 2000$. Statistically significant improvements are not easily obtained using poorly estimated ARFIMA models, and one may be better off in many cases using a simpler "approximation" to the true DGP, such as an ARIMA model, for the varieties of sample sizes available when using macroeconomic time series. Nevertheless, we have also seen, via our empirical analysis, that large samples or use of the prediction based estimator are sufficient to yield surprisingly strong support for ARFIMA models. In fact, even with very small samples, such as in our macroeconomic analysis, there is often a lack of evidence against ARFIMA models (as 130 of 215 variables had lower (or equal) point MSFEs for ARFIMA-best models, versus 85 "wins" for non-ARFIMA best models), and use of the prediction based estimator of d yields even more compelling evidence in favor of ARFIMA models. Given our Monte Carlo findings, such a high incidence of MSFE "wins" for the ARFIMA models in the SW dataset suggests that the sorts of models that may be prevalent are ones with d values less than 0.75. Inspection of Table 3 suggests that this may indeed be the case.

5   Concluding Remarks

We present the results of an empirical and Monte Carlo investigation of the usefulness of ARFIMA models in practical prediction based applications, and find strong evidence that such models may yield "better" approximations to unknown underlying DGPs than a fairly wide class of the benchmark type linear models, including AR, ARMA, ARIMA, and random walk models, among others. Furthermore, there appears little to choose between various estimators of d when samples are as large as often encountered in financial economics. For shorter samples such as those encountered in macroeconomics, we suggest the use of a prediction based estimator of d which is robust to finite sample bias and lag polynomial truncation effects that appear to plague standard in-sample estimators. We conclude that long memory processes, and in particular ARFIMA processes, might not fall into the "empty box" category after all, although much further research is needed before overwhelmingly conclusive evidence in either direction can be given.

6   References

Agiakloglou, C., P. Newbold and M. Wohar, 1992, Bias in an Estimator of the Fractional Difference Parameter, *Journal of Time Series Analysis*, 14, 235-246.

Andrews, D.W.K. and Y. Sun, 2002, Adaptive Local Whittle Estimation of Long-range Dependence, Working Paper, Yale University.

Baillie, R.T., 1996, Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics*, 73, 5-59.

Bank of Sweden, 2003, Time-Series Econometrics: Cointegration and Autoregressive Conditional Heteroskedasticity, *Advanced Information on the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel*, The Royal Swedish Academy of Sciences.

Beran, J., 1995, Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving Average Models, *J. R. Statist. Soc. B*, 57, No. 4, 659-672.

Bos, C.S., P.H. Franses and M. Ooms, 2002, Inflation, Forecast Intervals and Long Memory Regression Models, *International Journal of Forecasting*, 18, 243-264.

Breitung, Jörg and U. Hassler, 2002, Inference on the Cointegration Rank in Fractionally Integrated Processes, *Journal of Econometrics*, 110, 167-185.

Cheung, Y.-W., 1993, Tests for Fractional Integration: A Monte Carlo Investigation, *Journal of Time Series Analysis*, 14, 331-345.

Cheung, Y.-W. and F.X. Diebold, 1994, On Maximum Likelihood Estimation of the Difference Parameter of Fractionally Integrated Noise with Unknown Mean, *Journal of Econometrics*, 62, 301-316.

Chao, J.C., V. Corradi and N.R. Swanson, 2001, An Out of Sample Test for Granger Causality, *Macroeconomic Dynamics,* 5, 598-620.

Chio, K. and E. Zivot, 2002, Long Memory and Structural Changes in the Forward Discount: An Empirical Investigation, Working Paper, University of Washington.

Christoffersen, P.F., 1998, Evaluating Interval Forecasts, *International Economic Review*, 39, 841-862.

Christoffersen, P. and F.X. Diebold, 1997, Optimal Prediction Under Asymmetric Loss, *Econometric Theory, 13, 808-817.*

Clark, T.E. and M.W. McCracken, 2001, Tests of Equal Forecast Accuracy and Encompassing for Nested Models, *Journal of Econometrics,* 105, 85-110.

Clements, M.P. and J. Smith, 2000, Evaluating the Forecast Densities of Linear and Nonlinear Models: Applications to Output Growth and Unemployment, *Journal of Forecasting,* 19, 255-276.

Clements, M.P. and J. Smith, 2002, Evaluating Multivariate Forecast Densities: A Comparison of Two Approaches, *International Journal of Forecasting,* 18, 397-407.

Corradi, V. and N.R. Swanson, 2002, A Consistent Test for Out of Sample Nonlinear Predictive Ability, *Journal of Econometrics,* 110, 353-381.

Corradi, V. and N.R. Swanson, 2003, The Block Bootstrap for Parameter Estimation Error in Recursive Estimation Schemes, With Applications to Predictive Evaluation, Working Paper, Rutgers University.

Diebold, F.X., T. Gunther and A.S. Tay, 1998, Evaluating Density Forecasts with Applications to Finance and Management, *International Economic Review,* 39, 863-883.

Diebold, F.X., J. Hahn and A.S. Tay, 1999, Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Returns on Foreign Exchange, *Review of Economics and Statistics*, 81, 661-673.

Diebold, F. and A Inoue, 2001, Long Memory and Regime Switching, *Journal of Econometrics,* 105, 131-159.

Diebold, F.X. and R.S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X. and G.D. Rudebusch, 1989, Long Memory and Persistence in Aggregate Output, *Journal of Monetary Economics*, 24, 189-209.

Diebold, F.X. and G.D. Rudebusch, 1991a, Is Consumption Too Smooth? Long Memory and the Deaton Paradox, *Review of Economics and Statistics*, 73, 1-9.

Diebold, F.X. and G.D. Rudebusch, 1991b, On the Power of the Dickey-Fuller Test Against Fractional Alternatives, *Economics Letters*, 35, 155-160.

Ding, Z, C.W.J. Granger and R.F. Engle, 1993, A Long Memory Property of Stock Returns and a New Model, *Journal of Empirical Finance*, 1, 83-106.

Dittman, I. and C.W.J. Granger, 2002, Properties of Nonlinear Transformations of Fractionally Integrated Processes, *Journal of Econometrics*, 110, 113-133.

Engle, R.F. and A.D. Smith, 1999, Stochastic Permanent Breaks, *Review of Economics and Statistics*, 81, 553-574.

Geweke, J. and S. Porter-Hudak, 1983, The estimation and application of long memory time series models, *Journal of Time Series Analysis*, 4, 221-238.

Granger, C.W.J., 1969, Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica,* 37, 424-438.

Granger, C.W.J., 1980, Long Memory Relationships and the Aggregation of Dynamic Models, *Journal of Econometrics,* 14, 227-238.

Granger, C.W.J., 1999, Aspects of Research Strategies for Time Series Analysis, *Presentation to the conference on New Developments in Time Series Economics*, Yale University.

Granger, C.W.J. and A.P. Andersen, 1978, *Introduction to Bilinear Time Series Models,* Vandenhoeck and Ruprecht: Göttingen.

Granger, C.W.J., and Z. Ding, 1996, Varieties of Long Memory Models, *Journal of Econometrics,* 73, 61-77.

Granger, C.W.J. and M. Hatanaka, 1964, *Spectral Analysis of Economic Time Series,* Princeton University Press: Princeton.

Granger, C.W.J. and N. Hyung, 1999, Occasional Structural Breaks and Long Memory, Working Paper, University of California, San Diego.

Granger, C.W.J. and R. Joyeux, 1980, An Introduction to Long Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, 1, 15-30.

Granger, C.W.J. and P. Newbold, 1986, *Forecasting Economic Time Series*, Academic Press: San Diego.

Harvey, D.I., S.J. Leybourne and P. Newbold, (1997), Tests for Forecast Encompassing, *Journal of Business and Economic Statistics,* 16, 254-259.

Hassler, U. and J. Wolters, 1995, Long Memory in Inflation Rates: International Evidence, *Journal of Business and Economic Statistics*, 13, 37-45.

Hosking, J. 1981, Fractional Differencing, *Biometrica*, 68, 165-76.

Hurst, H.E., 1951, Long-term Storage Capacity of Resevoirs, *Transactions of the American Society of Civil Engineers*, 116, 770-799.

Hyung, N. and P.H. Franses, 2001, Structural Breaks and Long Memory in US Inflation Rates: Do They Matter for Forecasting?, Working Paper, Erasmus University.

Künsch, H.R., 1987, Statistical Aspects of Self-similar Processes, in *Proceedings of the first World Congress of the Bernoulli Society*, 1, 67-74, ed. by Y. Prohorov and V.V. Sasanov, Utrecht, VNU Science Press.

Lee, D. and P. Schmidt, 1996, On the Power of the KPSS Test of Stationarity Against Fractionally-Integrated Alternatives, *Journal of Econometrics*, 73, 285-302.

Leybourne, S., D. Harris and B. McCabe, 2003, A Robust Test for Short Memory, Working Paper, University of Nottingham.

Lo, A. 1991, Long-Term Memory in Stock Market Prices, *Econometrica*, 59, 1279-1313.

McCracken, M.W., 1999, Asymptotics for Out of Sample Tests of Causality, Working Paper, Louisiana State University.

Phillips, P.C.B., 1987, Time Series Regression with a Unit Root, *Econometrica*, 55, 277-301.

Robinson, P., 1995a, Log-Periodogram Regression of Time Series with Long Range Dependence, *The Annals of Statistics*, 23, 1048- 1072.

Robinson, P., 1995b, Gaussian Semiparametric Estimation of Long Range Dependence, *The Annals of Statistics*, 23, 1630- 1661.

Shimotsu, K. and P.C.B. Phillips, 2002, Exact Local Whittle Estimation of Fractional Integration, Working Paper, University of Essex.

Sowell, F.B., 1992a, Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, *Journal of Econometrics*, 53, 165-188.

Sowell, F.B., 1992b, Modelling Long-Run Behavior with the Fractional ARIMA Model, *Journal of Monetary Economics*, 29, 277-302.

Stock, J. and M. Watson, 2002, Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business and Economic Statistics*, 20, 147-162.

Taqqu, M. and V. Teverovsky, 1997, Robustness of Whittle-type Estimators for Time Series with Long-range Dependence, *Stochastic Models*, 13, 723-757.

van Dijk, D., P. Franses and R. Paap, 2002, A Nonlinear Long Memory odel, with an Application to US Unemployment, *Journal of Econometrics* 110, 135-165.

West, K., 1996, Asymptotic Inference About Predictive Ability, *Econometrica*, 64, 1067-1084.

White, H., 2000, A Reality Check for Data Snooping, *Econometrica*, 68, 1097-1126.

Table 1: Data Generating Processes for Monte Carlo Experiment[(∗)]

Name	DGP, ARFIMA(p,d,q)	Name	DGP, NON-ARFIMA
ARFIMA1	p=1, d=0.05, q=0, AR coef=0.60	RW	Random Walk without drift
ARFIMA2	p=1, d=0.05, q=1, AR coef=0.60, MA coef =-0.50	AR1	AR(p), p=1, AR coef=0.25
ARFIMA3	p=1, d=0.05, q=1, AR coef=0.60, MA coef =0.50	AR2	AR(p), p=1, AR coef=0.50
ARFIMA4	p=1, d=0.05, q=0, AR coef=0.75	AR3	AR(p), p=1, AR coef=0.75
ARFIMA5	p=1, d=0.05, q=1, AR coef=0.75, MA coef =-0.50	AR4	AR(p), p=1, AR coef=0.95
ARFIMA6	p=1, d=0.05, q=1, AR coef=0.75, MA coef =0.50	REG1	Regime Switching, Two States -1 and 1; Transition Probability 0.950
ARFIMA7	p=1, d=0.05, q=0, AR coef=0.95	REG2	Regime Switching, Two States -1 and 1; Transition Probability 0.990
ARFIMA8	p=1, d=0.05, q=1, AR coef=0.95, MA coef =-0.50	REG3	Regime Switching, Two States -1 and 1; Transition Probability 0.995
ARFIMA9	p=1, d=0.05, q=1, AR coef=0.95, MA coef =0.50	MA1	MA(q), q=1, MA coef= 0.75
ARFIMA10	p=1, d=0.25, q=0, AR coef=0.60	MA2	MA(q), q=1, AR MA coef= 0.25
ARFIMA11	p=1, d=0.25, q=1, AR coef=0.60, MA coef =-0.50	MA3	MA(q), q=1, MA coef= -0.25
ARFIMA12	p=1, d=0.25, q=1, AR coef=0.60, MA coef =0.50	MA4	MA(q), q=1, MA coef= -0.75
ARFIMA13	p=1, d=0.25, q=0, AR coef=0.75	GARCH1	AR-GARCH AR coef= 0.0 ARCH coef=0.2, GARCH Coef=0.2
ARFIMA14	p=1, d=0.25, q=1, AR coef=0.75, MA coef =-0.50	GARCH2	AR-GARCH AR coef= 0.5 ARCH coef=0.2, GARCH Coef=0.2
ARFIMA15	p=1, d=0.25, q=1, AR coef=0.75, MA coef =0.50	GARCH3	AR-GARCH AR coef= 0.9 ARCH coef=0.2, GARCH Coef=0.2
ARFIMA16	p=1, d=0.25, q=0, AR coef=0.95	ARX1	ARX, AR coef = 0.25, X coef = 0.50
ARFIMA17	p=1, d=0.25, q=1, AR coef=0.95, MA coef =-0.50	ARX2	ARX, AR coef = 0.50, X coef = 0.50
ARFIMA18	p=1, d=0.25, q=1, AR coef=0.95, MA coef =0.50	ARX3	ARX, AR coef = 0.75, X coef = 0.50
ARFIMA19	p=1, d=0.45, q=0, AR coef=0.60	ARX4	ARX, AR coef = 0.95, X coef = 0.50
ARFIMA20	p=1, d=0.45, q=1, AR coef=0.60, MA coef =-0.50	ARIMA1	ARIMA(p,d,q), p=1, d=0, q=1, AR coef= 0.25, MA coef= -0.5
ARFIMA21	p=1, d=0.45, q=1, AR coef=0.60, MA coef =0.50	ARIMA2	ARIMA(p,d,q), p=1, d=0, q=1, AR coef= 0.25, MA coef= 0.5
ARFIMA22	p=1, d=0.45, q=0, AR coef=0.75	ARIMA3	ARIMA(p,d,q), p=1, d=0, q=1, AR coef= 0.75, MA coef= -0.5
ARFIMA23	p=1, d=0.45, q=1, AR coef=0.75, MA coef =-0.50	ARIMA4	ARIMA(p,d,q), p=1, d=0, q=1, AR coef= 0.75, MA coef= 0.5
ARFIMA24	p=1, d=0.45, q=1, AR coef=0.75, MA coef =0.50	ARIMA5	ARIMA(p,d,q), p=1, d=1, q=1, AR coef= 0.25, MA coef= -0.5
ARFIMA25	p=1, d=0.45, q=0, AR coef=0.95	ARIMA6	ARIMA(p,d,q), p=1, d=1, q=1, AR coef= 0.25, MA coef= 0.5
ARFIMA26	p=1, d=0.45, q=1, AR coef=0.95, MA coef =-0.50	ARIMA7	ARIMA(p,d,q), p=1, d=1, q=1, AR coef= 0.75, MA coef= -0.5
ARFIMA27	p=1, d=0.45, q=1, AR coef=0.95, MA coef =0.50	ARIMA8	ARIMA(p,d,q), p=1, d=1, q=1, AR coef= 0.75, MA coef= 0.5
ARFIMA28	p=1, d=0.55, q=0, AR coef=0.60		
ARFIMA29	p=1, d=0.55, q=1, AR coef=0.60, MA coef =-0.50		
ARFIMA30	p=1, d=0.55, q=1, AR coef=0.60, MA coef =0.50		
ARFIMA31	p=1, d=0.55, q=0, AR coef=0.75		
ARFIMA32	p=1, d=0.55, q=1, AR coef=0.75, MA coef =-0.50		
ARFIMA33	p=1, d=0.55, q=1, AR coef=0.75, MA coef =0.50		
ARFIMA34	p=1, d=0.55, q=0, AR coef=0.95		
ARFIMA35	p=1, d=0.55, q=1, AR coef=0.95, MA coef =-0.50		
ARFIMA36	p=1, d=0.55, q=1, AR coef=0.95, MA coef =0.50		
ARFIMA37	p=1, d=0.75, q=0, AR coef=0.60		
ARFIMA38	p=1, d=0.75, q=1, AR coef=0.60, MA coef =-0.50		
ARFIMA39	p=1, d=0.75, q=1, AR coef=0.60, MA coef =-0.50		
ARFIMA40	p=1, d=0.75, q=0, AR coef=0.75		
ARFIMA41	p=1, d=0.75, q=1, AR coef=0.75, MA coef =-0.50		
ARFIMA42	p=1, d=0.75, q=1, AR coef=0.75, MA coef =0.50		
ARFIMA43	p=1, d=0.75, q=0, AR coef=0.95		
ARFIMA44	p=1, d=0.75, q=1, AR coef=0.95, MA coef =-0.50		
ARFIMA45	p=1, d=0.75, q=1, AR coef=0.95, MA coef =0.50		

[(∗)] Notes: This table summarizes all DGPs used the the Monte Carlo experiments reported in Section 4 of the paper.

Table 2: Analysis of Stock Market Data[*]

Index	Return Series	d Estimates	SM Rejec	MSFE best non-ARFIMA	MSFE best ARFIMA	DM Test
S&P500	$|r_t|$	(0.47, 0.33, 0.14, 0.47)	5	0.53,ARIMA(3,1,3)	0.52,RR*	-1.19,NONE
	r_t^2	(0.17, 0.14, 0.07, -0.18)	4	0.57,ARIMA(3,1,3)	0.57,GPH*	-0.94,NONE
	$\log\left(r_t^2\right)$	(0.40, -0.06, 0.12, 0.96)	5	0.50,ARIMA(0,1,1)	0.50,RR*	-0.12,NONE
	$|r_t|$, Pre 1987	(0.68, 0.56, 0.12, 0.55)	5	0.50,ARIMA(0,1,1)	0.50,GPH*	-0.11,NONE
	r_t^2, Pre 1987	(0.65, 0.51, 0.13, -0.30)	5	0.53,ARMA(1,1)	0.54,GPH	1.22,NONE
	$\log\left(r_t^2\right)$, Pre 1987	(0.85, 0.92, 0.09, 1.09)	5	0.49,ARIMA(0,1,1)	0.49,WHI*	-0.50,NONE
	$|r_t|$, Post 1987	(0.75, 1.11, 0.21, 0.48)	5	0.53,ARMA(1,1)	0.53,WHI*	0.10,NONE
	r_t^2, Post 1987	(0.58, 1.09, 0.16, -0.30)	4	0.59,ARMA(1,1)	0.59,RR*	0.00,NONE
	$\log\left(r_t^2\right)$, Post 1987	(0.70, 1.00, 0.20, 1.23)	4	0.50,ARMA(1,1)	0.50,WHI*	-0.35,NONE
FTSE	$|r_t|$	(0.38, 0.37, 0.14, 0.47)	4	0.53,ARMA(1,1)	0.54,GPH	3.56,ARMA
	r_t^2	(0.30, 0.21, 0.08, -0.16)	4	0.59,AR(5)	0.65,WHI	0.78,NONE
	$\log\left(r_t^2\right)$	(0.58, 1.15, 0.10, 1.03)	3	0.51,ARIMA(0,1,1)	0.51,RR*	-0.62,NONE
	$|r_t|$, Pre 1987	(0.74, 0.96, 0.19, 0.59)	4	0.56,ARMA(1,1)	0.58,GPH	3.30,ARMA
	r_t^2, Pre 1987	(0.30, 0.33, 0.12, -0.15)	3	1.00,RW	0.58,AML*	-4.94,AML
	$\log\left(r_t^2\right)$, Pre 1987	(0.83, 0.77, 0.15, 1.44)	4	0.55,MA(1)	0.56,GPH	0.91,NONE
	$|r_t|$, Post 1987	(0.49, -0.06, 0.10, 0.49)	4	0.53,ARMA(1,1)	0.53,RR*	-1.87, RR
	r_t^2, Post 1987	(0.36, 0.50, 0.07, -0.30)	3	0.54,ARMA(1,1)	0.54,RR*	0.02,NONE
	$\log\left(r_t^2\right)$, Post 1987	(0.51, 1.13, 0.12, 1.13)	4	0.51,ARIMA(0,1,1)	0.51,RR*	-0.23,NONE
DAX	$|r_t|$	(0.56, 0.51, 0.23, 0.50)	5	0.52,ARMA(1,1)	0.52,RR*	-2.36, RR
	r_t^2	(0.36, 0.30, 0.16, -0.15)	5	0.54,AR(9)	0.54,WHI*	0.43,NONE
	$\log\left(r_t^2\right)$	(0.64, 0.98, 0.22, 1.00)	5	0.50,ARMA(1,1)	0.50,GPH*	-0.26,NONE
	$|r_t|$, Pre 1987	(0.53, 1.12, 0.15, 0.54)	5	0.55,ARIMA(0,1,1)	0.54,WHI*	-1.23,NONE
	r_t^2, Pre 1987	(0.35, 0.20, 0.14, -0.37)	5	0.54,ARIMA(1,1,1)	0.54,GPH*	-0.53,NONE
	$\log\left(r_t^2\right)$, Pre 1987	(0.71, 1.14, 0.13, 1.43)	4	0.53,ARIMA(0,1,1)	0.53,RR*	-0.85,NONE
	$|r_t|$, Post 1987	(0.44, -0.49, 0.18, 0.52)	5	0.52,ARMA(1,1)	0.52,GPH*	1.43,ARMA
	r_t^2, Post 1987	(0.22, 0.44, 0.11, -0.17)	5	0.56,ARIMA(1,1,1)	0.56,WHI*	-1.81,WHI
	$\log\left(r_t^2\right)$, Post 1987	(0.59, 1.07, 0.18, 1.04)	5	0.51,ARMA(1,1)	0.51,RR*	-0.50,NONE
Nikkei 225	$|r_t|$	(0.52, 0.51, 0.22, 0.45)	5	0.55,ARIMA(1,1,2)	0.55,WHI*	1.69,ARIMA
	r_t^2	(0.20, 0.17, 0.12, -0.23)	5	0.57,ARIMA(2,1,3)	0.57,GPH*	0.53,NONE
	$\log\left(r_t^2\right)$	(0.73, 1.08, 0.21, 0.98)	5	0.50,ARIMA(1,1,1)	0.50,GPH*	-3.37,GPH
	$|r_t|$, Pre 1987	(0.54, 1.23, 0.19, 0.45)	5	0.58,ARIMA(1,1,2)	0.58,WHI*	0.36,NONE
	r_t^2, Pre 1987	(0.37, 0.13, 0.14, -0.48)	4	0.62,AR(1)	0.63,RR	1.11,NONE
	$\log\left(r_t^2\right)$, Pre 1987	(0.38, 0.08, 0.13, 1.14)	4	0.57,ARIMA(0,1,1)	0.56,WHI*	-0.75,NONE
	$|r_t|$, Post 1987	(0.60, 0.51, 0.26, 0.55)	5	0.53,ARMA(1,1)	0.53,RR*	0.71,NONE
	r_t^2, Post 1987	(0.46, 0.34, 0.20, -0.12)	4	0.53,ARMA(1,1)	0.53,WHI*	0.34,NONE
	$\log\left(r_t^2\right)$, Post 1987	(0.79, 0.55, 0.25, 1.06)	5	0.49,ARMA(1,1)	0.49,WHI*	1.19,NONE
Hang Sang	$|r_t|$	(0.47, 0.12, 0.19, 0.52)	5	0.58,ARMA(2,1)	0.58,GPH*	-0.69,NONE
	r_t^2	(0.01, 0.05, 0.02, 0.00)	0	0.73,ARX	0.73,RR*	-2.45, RR
	$\log\left(r_t^2\right)$	(0.66, 0.84, 0.21, 0.99)	4	0.52,ARMA(1,1)	0.52,WHI*	0.12,NONE
	$|r_t|$, Pre 1987	(0.46, 0.34, 0.11, 0.65)	4	0.54,ARIMA(2,1,2)	0.53,GPH*	-2.69,GPH
	r_t^2, Pre 1987	(0.35, 0.34, 0.10, 0.01)	3	0.59,ARMA(2,1)	0.54,AML*	-4.98,AML
	$\log\left(r_t^2\right)$, Pre 1987	(0.58, 1.16, 0.05, 1.10)	2	0.51,ARIMA(0,1,1)	0.51,RR*	0.46,NONE
	$|r_t|$, Post 1987	(0.47, 0.31, 0.17, 0.54)	5	0.57,ARIMA(2,1,3)	0.57,GPH*	-0.25,NONE
	r_t^2, Post 1987	(0.21, 0.19, 0.05, -0.08)	21	0.73,ARMA(1,1)	0.72,GPH*	-0.42,NONE
	$\log\left(r_t^2\right)$, Post 1987	(0.87, 1.16, 0.17, 1.11)	5	0.51,ARIMA(0,1,1)	0.51,WHI*	0.93,NONE

[*] Notes: Models are estimated as discussed above, and model acronyms used are outlined in Table 1. d estimates are based on GPH, AML, RR, and WHI estimators, respectively. SM Rejec reports the number of short memory test rejections based on application of all 5 SM tests discussed above, at a 5% nominal level. MFSEs are normalized by the RW MSFE. Starred entries correspond to cases for which the point MSFE of the ARFIMA-best model was lower (or the same) as for the non-ARFIMA-best model. All predictions are 1-step ahead, the out-of-sample period is equal to 50% of the sample, and Diebold and Mariano (DM) test is based on MSFE loss, assumes that PEE vanishes, and assmues the $N(0,1)$ distribution is asymptotically valid, as discussed in Section 2.3. Tests are based on nominal 20% level critical values. MSFE "best" non-ARFIMA (column 5) and ARFIMA (column 6) models are the ones compared with the DM test (column 7). DM test statistic values are reported, along with the "winning" model for those cases were the null of the DM test is rejected, and "NONE" when there is no significant difference between the models.

Table 3: Analysis of Macroeconomic Data[*]

Macroeconomic Variable	d Estimates	SM Rejec	MSFE best non-ARFIMA	MSFE best ARFIMA	DM Test
CCI30M	(1.16, 1.01, 0.32, 1.88)	5	0.92,ARIMA(1,1,1)	0.94,GPH	0.89,NONE
CCINRV	(1.11, 0.68, 0.27, 1.36)	4	0.67,ARIMA(1,1,1)	0.66,GPH*	-0.49,NONE
CCINT	(1.26, 0.99, 0.23, 0.92)	4	0.80,ARX	0.83,WHI	0.63,NONE
CCINV	(1.09, 0.98, 0.23, 0.97)	4	0.75,AR(3)	0.71,RR*	-1.26,NONE
CONDO9	(1.25, 1.13, 0.31, 2.27)	5	0.70,AR(3)	0.72,GPH	1.19,NONE
CONPC	(0.78, 0.51, 0.29, 2.11)	5	0.99,RW-D	1.01,WHI	0.61,NONE
CONQC	(0.78, 0.72, 0.30, 2.12)	5	1.00,RW	0.94,GPH*	-2.37,GPH
CONTC	(0.78, 0.51, 0.26, 2.12)	5	0.99,RW-D	1.03,WHI	1.53,RW-D
EXRCAN	(0.40, 0.29, 0.16, 0.09)	3	0.58,AR(1)	0.59,WHI	0.40,NONE
EXRGER	(0.27, 0.50, 0.08, 0.39)	1	0.66,MA(1)	0.66,RR*	-0.17,NONE
EXRJAN	(0.26, 0.20, 0.13, 0.16)	1	0.67,MA(1)	0.96,WHI	4.09,MA(1)
EXRSW	(0.46, 0.50, 0.16, 0.40)	2	0.64,MA(1)	0.66,RR	1.00,NONE
EXRUK	(0.19, 0.09, 0.11, -0.11)	0	0.52,AR(1)-GARCH	0.67,WHI	0.93,NONE
EXRUS	(0.09, 0.06, 0.07, 0.00)	0	0.65,MA(1)	0.65,AML*	-0.16,NONE
FCLBMC	(0.47, 0.50, 0.14, 0.33)	2	0.62,ARIMA(1,1,1)	0.62,WHI*	-0.87,NONE
FCLIN	(1.04, 1.05, 0.25, 0.66)	3	0.93,AR(1)	0.90,WHI*	-1.44,WHI
FCLNBF	(0.43, 0.17, 0.16, -0.47)	2	0.68,AR(1)	0.70,RR	0.46,NONE
FCLNQ	(0.67, 1.32, 0.17, 0.29)	3	0.61,ARIMA(1,1,1)	0.61,WHI*	-0.14,NONE
FCLRE	(0.93, 0.98, 0.21, 0.77)	3	0.90,ARIMA(1,1,1)	0.88,WHI*	-1.70,WHI
FCLS	(0.62, 1.39, 0.14, 0.73)	3	0.77,AR(1)	0.73,GPH*	-1.39,GPH
FCSGV	(0.68, 1.50, 0.15, 0.43)	2	0.74,ARX	0.68,WHI*	-3.35,WHI
FM1	(-1.35, -0.22, -0.18, -1.63)	2	0.31,ARX	0.79,RR	1.12,NONE
FM2	(-0.86, —, -0.07, -2.11)	2	0.41,MA(2)	0.43,GPH	1.44,MA(2)
FM2DQ	(0.92, 0.63, 0.19, 0.38)	3	0.86,AR(1)	0.85,AML*	-0.35,NONE
FM3	(-0.67, -0.44, -0.10, -1.21)	2	0.34,MA(1)	0.33,AML*	-1.99,AML
FMFBA	(-0.49, 0.08, -0.20, -1.13)	2	0.31,AR(5)	0.30,GPH*	-0.86,NONE
FML	(-0.38, 0.42, -0.16, -0.81)	2	0.29,ARIMA(0,1,1)	0.29,GPH*	0.20,NONE
FMRNBC	(-0.89, -0.49, -0.22, -0.90)	2	0.24,MA(1)	0.22,AML*	-1.17,NONE
FMRRA	(-1.17, -0.14, -0.25, -1.29)	2	0.28,AR(9)	0.47,RR	2.59,AR(9)
FSDXP	(1.23, 1.01, 0.31, 1.62)	4	0.97,ARIMA(1,1,1)	0.96,WHI*	-0.17,NONE
FSNCOM	(-0.14, -0.07, 0.05, -0.28)	1	0.65,MA(1)	0.65,RR*	-0.33,NONE
FSNFI	(-0.32, -0.25, 0.04, -0.94)	2	0.69,MA(1)	0.69,RR*	-0.74,NONE
FSNIN	(-0.31, -0.49, 0.05, -0.76)	1	0.70,MA(1)	0.69,RR*	-0.77,NONE
FSNTR	(-0.66, -0.10, 0.07, -1.16)	3	0.69,MA(1)	0.68,RR*	-0.46,NONE
FSNUT	(-0.12, -0.15, 0.05, -0.46)	2	0.57,AR(1)	1.42,AML	1.22,NONE
FSNVV3	(0.01, -0.48, -0.12, 0.75)	1	0.34,AR(2)	0.62,WHI	5.08,AR(2)
FSPCAP	(-0.22, -0.49, 0.01, -0.29)	0	0.62,MA(1)	0.61,WHI*	-0.41,NONE
FSPCOM	(-0.16, -0.06, 0.03, -0.28)	0	0.65,MA(1)	0.65,RR*	-1.05,NONE
FSPFI	(-0.19, -0.13, 0.04, -0.78)	1	0.65,AR(1)	0.65,RR*	0.05,NONE
FSPIN	(-0.16, -0.48, 0.03, -0.25)	0	0.66,MA(1)	0.65,RR*	-1.01,NONE
FSPTR	(-0.23, -0.48, 0.01, -0.27)	0	0.63,MA(1)	0.63,GPH*	-0.20,NONE
FSPUT	(-0.24, -0.11, 0.10, -0.48)	2	0.60,AR(1)	0.62,AML	1.53,NONE
FSPXE	(1.31, 0.51, 0.33, 1.73)	4	0.94,AR(2)	0.93,GPH*	-0.45,NONE
FSTB	(-0.24, -0.25, -0.12, -0.10)	0	0.38,MA(1)	0.41,WHI	1.73,MA(1)
FSTE	(0.69, —, -0.03, 0.77)	2	0.37,AR(1)	0.36,WHI*	-0.64,NONE
FSTM	(0.07, 0.04, -0.16, 0.55)	1	0.37,MA(2)	0.37,GPH*	-0.13,NONE
FTB	(-0.03, -0.27, -0.13, 0.03)	0	0.47,MA(1)	0.47,WHI*	-0.10,NONE
FTMD	(0.22, 0.01, -0.16, 0.55)	1	0.47,AR(1)	0.48,GPH	0.54,NONE
FWAFIT	(1.79, 2.38, 0.30, 2.02)	4	0.84,ARMA(2,1)	0.86,GPH	0.24,NONE
FYAAAC	(0.22, 0.04, 0.09, 0.20)	0	0.64,MA(1)	0.63,AML*	-1.35,AML
FYBAAC	(0.34, 0.19, 0.13, 0.01)	2	0.72,ARX	0.72,RR*	-0.15,NONE

[*] Notes: See notes to Table 2. Data used in this table correspond to those used in Stock and Watson (2002), and the variables are monthly, spanning the period 1959-1998 (see above for further details). For the following series we could not estimate AML due to convergence problems: FM2, FSTE, LEHCC, PUNEW, PUXHS, PUXM.

Table 3: Analysis of Macroeconomic Data (cont.)$^{(*)}$

Macroeconomic Variable	d Estimates	SM Rejec	MSFE best non-ARFIMA	MSFE best ARFIMA	DM Test
FYCP90	(-0.15, -0.10, 0.06, -0.48)	1	0.63,ARIMA(0,1,1)	0.64,RR	1.08,NONE
FYFF	(-0.12, -0.49, 0.07, -0.50)	1	0.68,AR(1)	0.68,GPH*	-0.34,NONE
FYFHA	(0.09, -0.11, 0.05, -0.04)	0	0.66,MA(2)	0.66,AML*	-0.36,NONE
FYGM3	(0.01, -0.49, 0.02, -0.33)	0	0.68,MA(1)	0.68,WHI*	0.30,NONE
FYGM6	(0.02, 0.03, 0.02, -0.32)	0	0.67,MA(1)	0.67,AML*	0.09,NONE
FYGT1	(-0.01, -0.03, 0.04, -0.24)	0	0.63,MA(1)	0.62,AML*	-1.28,AML
FYGT10	(0.17, -0.13, 0.01, 0.06)	0	0.64,MA(1)	0.64,WHI*	-0.27,NONE
FYGT5	(0.01, -0.13, 0.02, -0.09)	0	0.66,MA(1)	0.64,AML*	-0.50,NONE
GMCANQ	(-0.14, -0.15, -0.07, -0.12)	0	0.45,AR(1)	0.47,AML	0.53,NONE
GMCDQ	(0.29, 0.06, -0.05, 0.35)	1	0.40,MA(3)	0.41,RR	0.45,NONE
GMCNQ	(0.59, 1.07, -0.04, 0.34)	2	0.32,AR(1)	0.31,GPH*	-0.58,NONE
GMCQ	(0.82, 0.92, 0.01, 0.44)	2	0.40,MA(1)	0.37,WHI*	-2.00,WHI
GMCSQ	(0.59, 0.96, 0.04, 0.49)	2	0.40,AR(1)	0.36,GPH*	-2.50,GPH
GMDC	(-0.31, 0.41, -0.15, -0.85)	2	0.28,MA(1)	0.27,AML*	-1.07,NONE
GMDCD	(-0.76, 0.35, -0.16, -1.21)	2	0.22,ARIMA(0,1,1)	0.22,RR*	2.46,ARIMA(0,1,1)
GMDCN	(-0.22, 0.37, -0.14, -0.77)	2	0.38,AR(5)	0.39,GPH	0.28,NONE
GMDCS	(-0.31, -0.49, -0.20, -0.85)	1	0.18,ARIMA(0,1,1)	0.18,AML*	0.02,NONE
GMPYQ	(0.77, 1.07, 0.07, 0.45)	2	0.42,AR(1)	0.41,GPH*	-0.45,NONE
GMYXPQ	(0.88, 1.23, 0.12, 0.41)	3	0.62,AR(1)	0.59,WHI*	-2.09,WHI
HHSNTN	(1.22, 0.81, 0.34, 1.81)	4	0.98,AR(1)GARCH	2.13,GPH	5.96,AR(1)GARCH
HMOB	(1.16, 1.30, 0.33, 2.30)	5	1.00,RW	1.04,GPH	0.64,NONE
HNIV	(0.82, 0.65, 0.36, 1.92)	5	0.82,ARMA(3,1)	0.86,WHI	3.00,ARMA(3,1)
HNR	(0.88, 0.83, 0.30, 1.86)	5	0.97,AR(1)	0.94,AML*	-0.91,NONE
HNS	(0.80, 0.84, 0.29, 1.91)	5	0.99,ARX	0.96,AML*	-1.12,NONE
HNSMW	(0.98, 0.91, 0.32, 2.64)	5	0.73,ARX	0.74,GPH	1.08,NONE
HNSNE	(0.79, 1.37, 0.28, 2.61)	5	0.71,ARMA(1,1)	0.72,GPH	0.92,NONE
HNSSOU	(0.86, 0.80, 0.22, 2.66)	5	0.86,ARIMA(1,1,1)	0.84,GPH*	-0.75,NONE
HNSWST	(0.95, 0.94, 0.29, 2.64)	5	0.89,AR(2)	0.90,GPH	0.93,NONE
HSBMW	(0.81, 0.94, 0.25, 1.93)	4	0.94,ARMA(1,1)	0.94,GPH*	-0.04,NONE
HSBNE	(0.82, 1.03, 0.30, 1.94)	4	0.98,ARIMA(1,1,1)	0.94,WHI*	-0.91,NONE
HSBR	(0.89, 0.51, 0.24, 2.38)	4	0.98,AR(1)	1.14,GPH	1.69,AR(1)
HSBSOU	(0.80, 0.57, 0.27, 1.96)	4	0.98,AR(2)	1.01,GPH	0.82,NONE
HSBWST	(0.75, 0.57, 0.26, 1.95)	4	0.97,AR(1)	1.04,GPH	0.94,NONE
HSFR	(0.87, 0.69, 0.24, 2.38)	4	0.89,ARX	0.95,GPH	0.82,NONE
HSMW	(0.81, 1.50, 0.24, 2.33)	4	0.86,AR(4)	0.81,GPH*	-1.44,GPH
HSNE	(0.92, 1.23, 0.25, 2.38)	4	0.75,AR(1)-GARCH	0.71,GPH*	-1.07,NONE
HSSOU	(0.96, 1.16, 0.26, 2.34)	4	0.79,ARIMA(2,1,2)	0.81,GPH	0.58,NONE
HSWST	(0.97, 1.12, 0.27, 2.30)	4	0.85,ARMA(1,1)	0.87,GPH	1.88,ARMA(1,1)
IP	(0.42, 0.15, 0.08, 0.33)	1	0.66,AR(1)	0.65,RR*	-1.20,NONE
IPC	(0.18, -0.02, 0.04, 0.37)	1	0.46,AR(1)	0.46,GPH*	-0.50,NONE
IPCD	(0.07, -0.14, 0.01, 0.20)	1	0.49,AR(1)	0.50,RR	0.53,NONE
IPCN	(0.33, 0.04, 0.01, 0.41)	2	0.44,MA(1)	0.41,AML*	-1.24,NONE
IPD	(0.37, 0.17, 0.07, 0.23)	1	0.64,AR(1)	0.67,WHI	1.53,AR(1)
IPE	(0.65, 0.51, 0.10, 0.33)	2	0.49,ARIMA(1,1,1)	0.49,WHI*	0.66,NONE
IPF	(0.50, 0.27, 0.10, 0.39)	2	0.53,MA(1)	0.51,WHI*	-1.19,NONE
IPI	(0.42, 0.24, 0.08, 0.30)	1	0.51,MA(3)	0.51,WHI*	0.00,NONE
IPM	(0.19, 0.06, 0.04, 0.25)	1	0.67,AR(1)	0.67,RR*	-0.04,NONE
IPMD	(0.08, -0.02, 0.04, 0.13)	1	0.69,AR(1)	0.77,WHI	2.45,AR(1)
IPMFG	(0.37, 0.16, 0.09, 0.31)	1	0.65,AR(1)	0.64,RR*	-1.09,NONE
IPMIN	(0.09, -0.34, -0.04, 0.24)	1	0.49,AR(1)	0.57,WHI	1.23,NONE
IPMND	(0.02, -0.31, 0.09, 0.39)	1	0.56,ARX	0.58,WHI	0.94,NONE
IPN	(0.16, 0.18, 0.09, 0.39)	1	0.54,AR(2)	0.52,WHI*	-1.28,WHI
IPP	(0.51, 1.24, 0.10, 0.37)	2	0.56,MA(1)	0.54,WHI*	-1.19,NONE
IPUT	(0.55, 0.51, 0.03, 0.53)	2	0.41,MA(1)	0.38,WHI*	-2.14,WHI
IPX	(0.89, 0.51, 0.23, 1.89)	4	0.97,AR(1)GARCH	1.09,GPH	1.21,NONE
IPXDCA	(0.91, 0.53, 0.22, 1.87)	4	0.97,AR(1)GARCH	1.16,GPH	2.03,AR(1)GARCH
IPXMCA	(1.07, 0.57, 0.28, 2.24)	3	0.99,AR(2)	0.98,GPH*	-0.17,NONE
IPXMIN	(0.96, 0.86, 0.30, 1.90)	4	1.00,RW-D	1.03,GPH	0.73,NONE
IPXNCA	(0.87, 0.59, 0.25, 1.91)	4	0.98,AR(1)	1.07,GPH	1.69,AR(1)

30

Table 3: Analysis of Macroeconomic Data (cont.)[*]

Macroeconomic Variable	d Estimates	SM Rejec	MSFE best non-ARFIMA	MSFE best ARFIMA	DM Test
IPXUT	(0.92, 0.83, 0.35, 1.92)	4	0.96,AR(1)	0.96,WHI*	-0.42,NONE
IVMFDQ	(0.77, 1.28, 0.17, 0.43)	4	0.68,AR(3)	0.67,RR*	-0.25,NONE
IVMFGQ	(0.91, 1.29, 0.17, 0.45)	3	0.71,ARIMA(1,1,1)	0.69,RR*	-1.28,RR
IVMFNQ	(0.42, 0.46, 0.09, 0.45)	2	0.66,AR(1)	0.63,WHI*	-1.95,WHI
IVMTQ	(0.73, 1.30, 0.14, 0.50)	3	0.68,ARIMA(1,1,1)	0.66,WHI*	-0.73,NONE
IVRRQ	(0.38, -0.48, 0.08, 0.46)	3	0.60,AR(1)	0.60,RR*	-1.25,NONE
IVSRMQ	(0.16, 0.07, 0.03, -0.92)	1	0.43,AR(1)	0.42,GPH*	-0.34,NONE
IVSRQ	(0.21, 0.16, 0.02, -0.41)	1	0.43,AR(1)	0.43,AML*	-0.28,NONE
IVSRRQ	(-0.15, -0.22, -0.01, -0.43)	1	0.42,MA(1)	0.42,GPH*	-0.69,NONE
IVSRWQ	(0.11, 0.12, -0.02, 0.06)	0	0.39,MA(1)	0.39,AML*	0.04,NONE
IVWRQ	(0.59, 0.53, 0.05, 0.59)	2	0.56,AR(1)	0.54,RR*	-2.20,RR
LEH	(-0.76, -0.49, -0.19, -1.32)	4	0.11,AR(8)	0.15,AML	3.86,AR(8)
LEHCC	(-0.91, —, -0.20, -1.28)	2	0.09,ARIMA(0,1,2)	0.09,RR*	0.61,NONE
LEHFR	(-0.92, 0.10, -0.27, -1.73)	3	0.06,AR(8)	0.12,AML	5.32,AR(8)
LEHM	(-0.74, -0.05, -0.17, -1.17)	2	0.14,AR(9)	0.17,GPH	2.54,AR(9)
LEHS	(-0.80, 0.07, -0.24, -0.95)	3	0.10,AR(5)	0.13,RR	2.77,AR(5)
LEHTT	(-1.20, -0.10, -0.28, -1.66)	4	0.12,ARX	0.17,RR	1.84,ARX
LEHTU	(-0.58, -0.11, -0.21, -0.92)	3	0.12,ARX	0.15,RR	3.93,ARX
LHEL	(0.43, 0.50, 0.13, -0.15)	2	0.43,AR(3)	0.44,WHI	0.95,NONE
LHELX	(1.17, 0.65, 0.32, 0.34)	4	0.85,ARMA(2,1)	0.90,GPH	1.54,ARMA(2,1)
LHEM	(0.58, 0.70, 0.03, 0.28)	1	0.55,MA(1)	0.53,WHI*	-1.04,NONE
LHNAG	(0.50, 0.99, 0.04, 0.31)	2	0.56,AR(1)	0.52,WHI*	-1.32,WHI
LHU14	(1.40, 1.13, 0.33, 1.61)	4	1.00,RW	0.98,GPH*	-0.99,NONE
LHU15	(1.54, 1.64, 0.30, 1.33)	4	0.90,AR(p)	0.90,GPH*	0.01,NONE
LHU26	(1.25, 1.25, 0.29, 1.37)	3	1.00,AR(1)-GARCH	1.03,GPH	0.82,NONE
LHU5	(1.30, 0.94, 0.35, 1.80)	4	0.81,ARIMA(2,1,0)	0.81,GPH*	0.07,NONE
LHU680	(1.35, 1.35, 0.34, 1.82)	4	0.90,ARMA(2,2)	0.92,GPH	0.71,NONE
LHUR	(1.55, 1.51, 0.31, 1.74)	4	1.00,RW-D	0.99,GPH*	-0.14,NONE
LP	(0.62, 0.53, 0.11, 0.28)	1	0.64,AR(2)	0.62,GPH*	-0.85,NONE
LPCC	(0.22, 0.26, 0.01, -0.01)	1	0.69,MA(1)	0.60,AML*	-1.61,AML
LPED	(0.40, 0.27, 0.11, -0.12)	1	0.82,ARIMA(1,1,1)	0.79,GPH*	-0.74,NONE
LPEM	(0.41, -0.49, 0.13, -0.13)	2	0.84,AR(2)	0.83,RR*	-1.05,NONE
LPEN	(0.28, 0.27, 0.16, -0.19)	1	0.69,ARIMA(2,1,0)	0.69,RR*	0.55,NONE
LPFR	(0.94, 0.89, 0.14, 0.41)	4	0.66,AR(5)	0.68,RR	1.09,NONE
LPGD	(0.40, -0.49, 0.12, -0.09)	2	0.79,ARIMA(1,1,1)	0.81,GPH	0.29,NONE
LPGOV	(0.91, 1.21, 0.14, 0.43)	4	0.61,ARIMA(3,1,1)	0.58,GPH*	-2.00,GPH
LPHRM	(0.83, 0.86, 0.28, 2.54)	3	0.88,ARMA(1,1)	0.84,GPH*	-1.82,GPH
LPMI	(-0.11, -0.15, 0.02, 0.11)	1	0.70,AR(1)	0.79,GPH	2.52,AR(1)
LPMOSA	(1.30, 1.13, 0.28, 1.78)	4	0.99,AR(2)	0.95,GPH*	-1.18,NONE
LPNAG	(0.79, 0.51, 0.11, 0.32)	1	0.65,AR(2)	0.61,GPH*	-0.61,NONE
LPS	(0.83, 1.12, 0.09, 0.49)	2	0.64,ARIMA(1,1,1)	0.63,WHI*	-0.50,NONE
LPSP	(1.01, 0.52, 0.13, 0.41)	3	0.44,ARIMA(1,1,2)	0.44, RR*	-0.61,NONE
LPT	(0.98, 1.28, 0.10, 0.38)	3	0.86,ARIMA(1,1,2)	0.82,WHI*	-1.43,WHI
LPTU	(0.47, 0.19, 0.03, 0.06)	1	0.32,AR(1)	0.32,WHI*	-1.24,NONE
LW	(0.21, -0.09, -0.04, 0.29)	1	0.26,MA(1)	0.26,GPH*	-2.19,GPH
MDO	(0.25, 0.16, 0.01, 0.45)	1	0.36,AR(1)	0.36,AML*	0.14,NONE
MDOQ	(0.17, 0.07, 0.01, 0.04)	0	0.36,AR(1)	0.37,WHI	0.70,NONE
MDU	(0.62, 0.72, 0.14, 0.44)	4	0.76,ARIMA(1,1,1)	0.76,RR*	-1.14,NONE
MDUWU	(0.22, 0.22, -0.01, 0.42)	1	0.32,MA(1)	0.33,GPH	0.60,NONE
MNO	(0.59, 0.95, 0.11, 0.63)	2	0.41,ARIMA(3,1,3)	0.41,GPH*	0.60,NONE
MNOU	(0.43, 0.14, 0.01, 0.57)	2	0.29,MA(1)	0.29,AML*	-0.90,NONE
MNU	(0.05, 0.29, 0.10, 0.26)	0	0.64,ARIMA(2,1,2)	0.65,GPH	2.69,ARIMA(2,1,2)
MO	(0.41, 0.11, 0.06, 0.56)	2	0.39,AR(1)	0.40,AML	0.27,NONE
MOCMQ	(0.17, -0.03, 0.02, 0.02)	0	0.46,MA(1)	0.46,WHI*	-0.19,NONE
MOWU	(0.27, 0.12, 0.01, 0.46)	1	0.34,AR(1)	0.35,AML	1.20,NONE
MPCON	(0.25, -0.14, -0.12, 0.52)	1	0.27,MA(1)	0.27,GPH*	0.81,NONE
MPCONQ	(0.12, -0.01, -0.09, 0.21)	0	0.27,MA(1)	0.27,GPH*	-0.68,NONE
MSDQ	(0.42, -0.05, 0.01, 0.22)	1	0.40,MA(1)	0.40,WHI*	-0.40,NONE
MSMQ	(0.48, 0.02, 0.02, 0.27)	1	0.40,MA(1)	0.39,WHI*	-0.52,NONE

Table 3: Analysis of Macroeconomic Data (cont.)[(*)]

Macroeconomic Variable	d Estimates	SM Rejec	MSFE best non-ARFIMA	MSFE best ARFIMA	DM Test
MSMTQ	(0.38, 0.10, 0.03, 0.36)	0	0.42,AR(1)	0.41,WHI*	-0.83,NONE
MSNQ	(0.23, 0.01, 0.01, 0.34)	0	0.40,AR(1)	0.39,AML*	-0.67,NONE
MSONDQ	(0.21, -0.01, -0.09, 0.18)	0	0.26,ARIMA(0,1,1)	0.26,WHI*	0.38,NONE
MU	(0.61, 0.77, 0.14, 0.44)	4	0.77,ARIMA(1,1,1)	0.77,RR*	-0.76,NONE
PCGOLD	(-0.86, -0.49, -0.11, -1.18)	3	0.18,ARIMA(0,1,2)	0.19,RR	2.34,ARIMA(0,1,2)
PMCP	(1.24, 0.51, 0.27, 1.83)	3	1.00,RW	1.25,RR	2.00,RW
PMDEL	(1.22, 1.21, 0.23, 1.35)	3	0.98,ARX	1.01,GPH	0.63,NONE
PMEMP	(1.13, 1.17, 0.16, 1.76)	3	0.98,AR(1)-GARCH	0.99,GPH	0.18,NONE
PMI	(0.99, 0.51, 0.14, 1.80)	3	0.96,AR(2)	0.98,GPH	0.38,NONE
PMNO	(1.01, 0.83, 0.17, 1.82)	3	0.96,ARX	1.01,AML	0.63,NONE
PMNV	(0.99, 0.82, 0.16, 1.87)	3	0.96,ARX	1.05,GPH	2.17,ARX
PMP	(1.01, 0.81, 0.17, 1.80)	3	0.97,ARX	1.03,GPH	0.71,NONE
PSM99Q	(-0.77, 0.36, -0.12, -1.56)	2	0.40,AR(3)	0.37,AML*	-1.81,AML
PU81	(-0.69, -0.49, -0.10, -1.72)	3	0.23,ARIMA(0,1,1)	0.23,RR*	2.07,ARIMA(0,1,1)
PU83	(-0.49, 0.23, -0.12, -0.85)	3	0.19,MA(1)	0.20,RR	0.89,NONE
PU84	(-0.79, 0.22, -0.19, -1.08)	3	0.47,MA(1)	0.43,AML*	-2.17,AML
PU85	(-0.35, 0.40, -0.19, -0.9)	2	0.25,AR(1)	0.26,RR	0.17,NONE
PUC	(-0.38, -0.47, -0.14, -0.82)	2	0.39,ARIMA(0,1,1)	0.41,RR	2.21,ARIMA(0,1,1)
PUCD	(-0.62, -0.49, -0.19, -0.77)	2	0.37,MA(1)	0.33,AML*	-2.53,AML
PUH	(-0.64, -0.49, -0.12, -1.1)	3	0.17,MA(1)	0.17,RR*	2.62,MA(1)
PUNEW	(-0.07, —, -0.15, -0.78)	2	0.38,ARIMA(0,1,1)	0.39,GPH	2.65,ARIMA(0,1,1)
PUS	(-0.47, -0.42, -0.26, -1.05)	2	0.30,MA(3)	0.31,AML	0.24,NONE
PUXF	(-0.14, -0.30, -0.21, -0.91)	1	0.33,ARIMA(0,1,1)	0.34,GPH	2.02,ARIMA(0,1,1)
PUXHS	(-0.26, —, -0.18, -0.80)	2	0.37,ARIMA(0,1,1)	0.39,RR	2.41,ARIMA(0,1,1)
PUXM	(-0.20, —, -0.17, -0.85)	2	0.36,MA(1)	0.38,AML	1.04,NONE
PW150A	(-1.02, -0.49, -0.16, -1.89)	4	0.32,ARIMA(1,1,1)	0.31,RR*	-0.59,NONE
PW160A	(-1.00, 0.21, -0.17, -1.24)	3	0.21,ARIMA(0,1,2)	0.24,RR	2.70,ARIMA(0,1,2)
PWCMSA	(-0.4, -0.02, -0.12, -0.54)	2	0.31,AR(5)	0.26,RR*	-2.94, RR
PWFCSA	(-0.46, 0.09, -0.17, -0.75)	1	0.30,ARIMA(0,1,1)	0.32,AML	1.37,ARIMA(0,1,1)
PWFSA	(-0.45, -0.49, -0.17, -0.80)	1	0.28,ARIMA(0,1,1)	0.29,AML	0.76,NONE
PWFXSA	(-0.34, -0.43, -0.09, -0.88)	4	0.29,ARIMA(0,1,1)	0.29,RR*	0.38,NONE
PWIMSA	(-0.78, -0.49, -0.09, -0.99)	2	0.47,AR(1)	0.52,AML	0.86,NONE
RTNQ	(0.57, 1.18, -0.04, 0.39)	2	0.33,AR(1)	0.35,WHI	1.01,NONE
RTQ	(0.48, -0.01, -0.01, 0.36)	1	0.38,MA(1)	0.38,GPH*	-0.70,NONE
SFYAAAC	(0.79, 0.54, 0.23, 0.38)	3	0.98,AR(1)	0.96,WHI*	-0.30,NONE
SFYBAAC	(0.90, 0.58, 0.23, 0.65)	3	0.98,AR(1)	1.00,WHI	0.17,NONE
SFYFHA	(1.03, 0.51, 0.26, 0.83)	3	0.90,ARX	0.92,WHI	0.99,NONE
SFYGM3	(0.84, 0.87, 0.27, 0.45)	3	0.93,AR(1)	0.93,AML*	-0.14,NONE
SFYGM6	(0.72, 0.86, 0.27, 0.08)	2	0.93,AR(1)	0.91,AML*	-0.59,NONE
SFYGT1	(0.73, 0.90, 0.27, -0.21)	2	0.94,AR(1)	0.90,AML*	-1.37,AML
SFYGT10	(0.77, 0.51, 0.25, 0.10)	2	0.97,AR(1)	0.97,RR*	-0.16,NONE
SFYGT5	(0.80, 0.51, 0.25, 0.09)	2	0.97,AR(1)	0.97,RR*	0.04,NONE
WTDQ	(0.37, 0.25, 0.03, 0.38)	1	0.40,AR(1)	0.40,AML*	0.12,NONE
WTNQ	(0.14, -0.20, -0.07, 0.41)	1	0.43,AR(1)	0.43,RR*	0.21,NONE
WTQ	(0.18, 0.08, -0.01, 0.42)	1	0.42,MA(1)	0.42,GPH*	-0.11,NONE
SFYCP90	(0.82, 0.98, 0.26, 0.30)	2	0.91,AR(1)	0.88,GPH*	-0.37,NONE

Table 4: Power of Tests of Short Memory with ARFIMA DGPs[*]

Model	d		Sample Size 500					Sample Size 2000			
		LHM	RR	MRR	GPH	WHI	LHM	RR	MRR	GPH	WHI
ARFIMA1	0.05	0.05	0.89	0.04	0.17	0.23	0.07	0.96	0.14	0.13	0.28
ARFIMA2		0.05	0.32	0.05	0.12	0.22	0.05	0.44	0.09	0.12	0.19
ARFIMA3		0.05	0.94	0.08	0.14	0.26	0.06	0.97	0.09	0.13	0.21
ARFIMA4		0.06	0.98	0.10	0.33	0.31	0.07	1.00	0.12	0.24	0.38
ARFIMA5		0.05	0.80	0.08	0.21	0.26	0.07	0.93	0.12	0.21	0.37
ARFIMA6		0.07	0.99	0.07	0.37	0.30	0.10	1.00	0.12	0.24	0.38
ARFIMA7		0.74	1.00	0.52	0.98	0.95	0.85	1.00	0.64	0.99	1.00
ARFIMA8		0.74	1.00	0.51	0.94	0.94	0.82	1.00	0.63	0.98	1.00
ARFIMA9		0.73	1.00	0.55	0.97	0.92	0.83	1.00	0.66	0.99	1.00
ARFIMA10	0.25	0.23	0.99	0.29	0.70	0.71	0.57	1.00	0.64	0.86	0.97
ARFIMA11		0.23	0.98	0.28	0.68	0.68	0.52	1.00	0.66	0.85	0.98
ARFIMA12		0.23	0.99	0.30	0.70	0.68	0.57	1.00	0.69	0.85	0.96
ARFIMA13		0.28	1.00	0.31	0.83	0.73	0.63	1.00	0.68	0.84	0.99
ARFIMA14		0.28	1.00	0.34	0.81	0.72	0.59	1.00	0.71	0.88	0.99
ARFIMA15		0.26	0.99	0.33	0.81	0.72	0.66	1.00	0.72	0.92	0.99
ARFIMA16		0.96	1.00	0.77	1.00	0.99	0.99	1.00	0.91	1.00	1.00
ARFIMA17		0.92	1.00	0.73	1.00	1.00	0.99	1.00	0.92	1.00	1.00
ARFIMA18		0.93	1.00	0.77	0.99	0.99	1.00	1.00	0.92	1.00	1.00
ARFIMA19	0.45	0.65	1.00	0.61	0.95	0.95	0.98	1.00	0.96	1.00	1.00
ARFIMA20		0.65	1.00	0.64	0.93	0.96	0.96	1.00	0.94	1.00	1.00
ARFIMA21		0.65	1.00	0.62	0.95	0.95	0.97	1.00	0.94	1.00	1.00
ARFIMA22		0.69	1.00	0.64	0.97	0.94	0.99	1.00	0.96	0.99	1.00
ARFIMA23		0.71	1.00	0.64	0.97	0.95	0.99	1.00	0.94	1.00	1.00
ARFIMA24		0.74	1.00	0.65	0.97	0.94	0.97	1.00	0.94	1.00	1.00
ARFIMA25		0.99	1.00	0.88	1.00	1.00	1.00	1.00	0.98	1.00	1.00
ARFIMA26		0.98	1.00	0.86	1.00	1.00	1.00	1.00	0.98	1.00	1.00
ARFIMA27		0.99	1.00	0.88	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA28	0.55	0.84	1.00	0.74	0.98	0.98	0.99	1.00	0.98	1.00	1.00
ARFIMA29		0.83	1.00	0.77	0.97	0.98	0.99	1.00	0.98	1.00	1.00
ARFIMA30		0.82	1.00	0.76	0.98	0.99	0.99	1.00	0.97	1.00	1.00
ARFIMA31		0.88	1.00	0.82	0.99	0.99	0.99	1.00	0.98	1.00	1.00
ARFIMA32		0.87	1.00	0.75	0.98	0.98	1.00	1.00	0.98	1.00	1.00
ARFIMA33		0.86	1.00	0.77	0.98	0.98	1.00	1.00	0.98	1.00	1.00
ARFIMA34		1.00	1.00	0.92	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA35		1.00	1.00	0.91	0.99	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA36		1.00	1.00	0.90	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA37	0.75	0.97	1.00	0.89	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA38		0.96	1.00	0.90	0.99	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA39		0.97	1.00	0.90	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA40		0.98	1.00	0.92	0.99	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA41		0.97	1.00	0.92	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ARFIMA42		0.99	1.00	0.89	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA43		1.00	1.00	0.95	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA44		1.00	1.00	0.97	1.00	1.00	1.00	1.00	0.99	1.00	1.00
ARFIMA45		1.00	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00

[*] Notes: Table reports the power of short memory tests for various ARFIMA models, using 5% nominal level critical values. Model acronyms used are outlined in Table 1. The five tests considered were LHM, Leybourne, Harris and McCabe (2003); RR, Rescaled Range; MRR, Modified Rescaled Range, Lo (1991); GPH, Geweke and Porter-Hudak (1983) and WHI, Local Whittle, Robinson (1995). LHM is based on the rate of decay of autocovariance function. RR and MRR are range based tests, while GPH and WHI are tests based on the estimated value of differencing paremeter d. Results are based on 500 Monte Carlo iterations.

Table 5: Size of Tests of Short Memory with Stationary Non-ARFIMA DGPs[*]

	Sample Size 500					Sample Size 2000				
Model	LHM	RR	MRR	GPH	WHI	LHM	RR	MRR	GPH	WHI
AR1	0.05	0.18	0.05	0.11	0.25	0.03	0.19	0.06	0.10	0.11
AR2	0.03	0.53	0.04	0.11	0.26	0.04	0.64	0.06	0.08	0.11
AR3	0.03	0.95	0.06	0.19	0.24	0.04	0.98	0.14	0.10	0.16
AR4	0.66	1.00	0.67	0.95	0.89	0.72	1.00	0.87	0.99	1.00
REG1	0.15	0.99	0.40	0.70	0.45	0.11	0.99	0.52	0.60	0.86
REG2	0.87	1.00	0.88	0.97	0.99	0.98	1.00	0.99	1.00	1.00
REG3	0.87	0.98	0.86	0.93	0.97	0.89	0.89	0.90	0.96	1.00
MA1	0.04	0.31	0.04	0.11	0.23	0.04	0.31	0.04	0.12	0.13
MA2	0.05	0.11	0.05	0.10	0.24	0.05	0.12	0.03	0.12	0.12
MA3	0.04	0.35	0.05	0.13	0.21	0.03	0.35	0.07	0.08	0.11
MA4	0.04	1.00	0.44	0.28	0.27	0.06	1.00	0.77	0.16	0.21
GARCH1	0.04	0.08	0.06	0.13	0.24	0.06	0.06	0.05	0.11	0.12
GARCH2	0.05	0.62	0.04	0.11	0.28	0.04	0.63	0.04	0.10	0.11
GARCH3	0.13	1.00	0.32	0.77	0.84	0.10	1.00	0.55	0.66	0.92
ARX1	0.05	0.15	0.03	0.11	0.22	0.06	0.24	0.04	0.11	0.11
ARX2	0.04	0.58	0.06	0.10	0.22	0.03	0.66	0.06	0.10	0.10
ARX3	0.03	0.97	0.10	0.23	0.25	0.04	0.98	0.14	0.12	0.18
ARX4	0.65	1.00	0.63	0.95	0.92	0.72	1.00	0.87	0.98	1.00
ARIMA1	0.02	0.53	0.07	0.12	0.23	0.07	0.50	0.08	0.10	0.11
ARIMA2	0.04	0.46	0.05	0.10	0.24	0.04	0.48	0.04	0.09	0.12
ARIMA3	0.04	0.64	0.07	0.17	0.24	0.04	0.74	0.07	0.09	0.14
ARIMA4	0.04	0.98	0.10	0.22	0.22	0.05	0.98	0.15	0.11	0.15

[*] Notes: See notes to Table 4. This table report the size properties of the short memory tests for all the stationary non-ARFIMA DGPs.

Table 6: Estimated d for ARFIMA DGPs$^{(*)}$

		Sample Size 500				Sample Size 2000			
Model	d	GPH	RR	WHI	AML	GPH	RR	WHI	AML
ARFIMA1	0.05	0.08(0.18)	0.15(0.04)	0.04(0.21)	-0.06(0.23)	0.06(0.10)	0.14(0.03)	0.07(0.08)	0.04(0.12)
ARFIMA2		0.04(0.17)	0.08(0.04)	0.03(0.21)	0.02(0.25)	0.05(0.10)	0.08(0.03)	0.06(0.07)	0.08(0.14)
ARFIMA3		0.08(0.17)	0.16(0.04)	0.04(0.21)	-0.10(0.23)	0.04(0.10)	0.14(0.03)	0.06(0.07)	0.02(0.11)
ARFIMA4		0.18(0.18)	0.18(0.04)	0.09(0.23)	-0.02(0.22)	0.09(0.11)	0.17(0.03)	0.10(0.08)	0.02(0.13)
ARFIMA5		0.12(0.18)	0.13(0.04)	0.05(0.21)	-0.02(0.38)	0.09(0.09)	0.13(0.03)	0.09(0.08)	0.03(0.27)
ARFIMA6		0.19(0.18)	0.19(0.04)	0.09(0.22)	-0.05(0.24)	0.09(0.11)	0.17(0.03)	0.10(0.08)	0.01(0.12)
ARFIMA7		0.72(0.16)	0.28(0.03)	0.58(0.20)	0.77(0.33)	0.49(0.10)	0.26(0.03)	0.54(0.08)	0.52(0.52)
ARFIMA8		0.68(0.18)	0.27(0.04)	0.57(0.21)	0.61(0.34)	0.48(0.11)	0.25(0.03)	0.52(0.08)	0.34(0.48)
ARFIMA9		0.73(0.18)	0.28(0.03)	0.56(0.21)	0.71(0.33)	0.48(0.11)	0.26(0.03)	0.53(0.08)	0.48(0.51)
ARFIMA10	0.25	0.35(0.19)	0.22(0.04)	0.35(0.22)	0.20(0.36)	0.29(0.11)	0.22(0.04)	0.30(0.08)	0.20(0.18)
ARFIMA11		0.32(0.19)	0.18(0.04)	0.33(0.21)	0.23(0.53)	0.28(0.10)	0.19(0.04)	0.29(0.08)	0.26(0.23)
ARFIMA12		0.35(0.17)	0.22(0.04)	0.33(0.21)	0.15(0.40)	0.28(0.12)	0.23(0.04)	0.29(0.09)	0.20(0.19)
ARFIMA13		0.44(0.18)	0.24(0.04)	0.37(0.22)	0.29(0.35)	0.31(0.10)	0.25(0.04)	0.33(0.08)	0.20(0.21)
ARFIMA14		0.42(0.18)	0.22(0.04)	0.38(0.23)	0.17(0.51)	0.31(0.11)	0.23(0.04)	0.32(0.07)	0.01(0.34)
ARFIMA15		0.44(0.18)	0.24(0.04)	0.36(0.22)	0.26(0.38)	0.33(0.11)	0.25(0.04)	0.34(0.08)	0.21(0.24)
ARFIMA16		0.97(0.18)	0.31(0.03)	0.85(0.21)	0.87(0.34)	0.72(0.10)	0.30(0.03)	0.75(0.08)	1.02(0.37)
ARFIMA17		0.93(0.18)	0.30(0.03)	0.85(0.20)	0.84(0.26)	0.71(0.10)	0.30(0.03)	0.74(0.08)	0.80(0.37)
ARFIMA18		0.97(0.19)	0.31(0.03)	0.84(0.22)	0.78(0.35)	0.72(0.10)	0.30(0.03)	0.75(0.08)	1.00(0.37)
ARFIMA19	0.45	0.61(0.18)	0.27(0.04)	0.64(0.23)	0.58(0.38)	0.53(0.10)	0.30(0.03)	0.54(0.08)	0.54(0.45)
ARFIMA20		0.57(0.17)	0.26(0.04)	0.63(0.21)	0.51(0.56)	0.50(0.11)	0.28(0.04)	0.52(0.08)	0.36(0.62)
ARFIMA21		0.61(0.17)	0.28(0.04)	0.62(0.22)	0.56(0.38)	0.51(0.10)	0.30(0.04)	0.53(0.08)	0.59(0.49)
ARFIMA22		0.69(0.19)	0.28(0.04)	0.65(0.24)	0.62(0.34)	0.56(0.11)	0.31(0.03)	0.57(0.08)	0.60(0.41)
ARFIMA23		0.67(0.18)	0.28(0.04)	0.66(0.22)	0.60(0.37)	0.54(0.11)	0.30(0.04)	0.55(0.08)	0.53(0.46)
ARFIMA24		0.71(0.18)	0.29(0.04)	0.66(0.24)	0.59(0.32)	0.56(0.10)	0.31(0.03)	0.56(0.08)	0.56(0.38)
ARFIMA25		1.21(0.17)	0.33(0.03)	1.12(0.20)	0.82(0.39)	0.94(0.11)	0.33(0.03)	0.97(0.08)	0.85(0.43)
ARFIMA26		1.15(0.19)	0.32(0.03)	1.09(0.20)	0.97(0.31)	0.93(0.10)	0.33(0.03)	0.96(0.08)	0.78(0.38)
ARFIMA27		1.18(0.19)	0.33(0.03)	1.11(0.20)	0.76(0.35)	0.94(0.11)	0.33(0.03)	0.97(0.08)	0.85(0.43)
ARFIMA28	0.55	0.73(0.18)	0.30(0.04)	0.77(0.22)	0.70(0.35)	0.62(0.11)	0.32(0.03)	0.64(0.08)	0.70(0.39)
ARFIMA29		0.71(0.18)	0.29(0.04)	0.75(0.22)	0.68(0.39)	0.61(0.10)	0.31(0.03)	0.63(0.08)	0.59(0.39)
ARFIMA30		0.75(0.18)	0.30(0.04)	0.77(0.22)	0.70(0.33)	0.63(0.11)	0.32(0.03)	0.65(0.08)	0.68(0.39)
ARFIMA31		0.84(0.17)	0.31(0.03)	0.81(0.23)	0.70(0.30)	0.65(0.11)	0.32(0.03)	0.67(0.08)	0.67(0.39)
ARFIMA32		0.80(0.19)	0.30(0.04)	0.80(0.22)	0.75(0.29)	0.65(0.10)	0.32(0.03)	0.66(0.08)	0.70(0.34)
ARFIMA33		0.83(0.18)	0.31(0.04)	0.79(0.22)	0.65(0.29)	0.66(0.11)	0.33(0.03)	0.68(0.08)	0.63(0.33)
ARFIMA34		1.26(0.17)	0.34(0.03)	1.21(0.19)	0.76(0.33)	1.04(0.11)	0.34(0.03)	1.08(0.08)	0.68(0.28)
ARFIMA35		1.25(0.18)	0.33(0.03)	1.22(0.20)	1.03(0.34)	1.03(0.10)	0.34(0.03)	1.07(0.08)	0.77(0.37)
ARFIMA36		1.28(0.18)	0.33(0.03)	1.23(0.19)	0.84(0.37)	1.04(0.11)	0.34(0.03)	1.08(0.08)	0.66(0.27)
ARFIMA37	0.75	0.99(0.17)	0.33(0.03)	1.03(0.21)	0.85(0.30)	0.83(0.11)	0.34(0.03)	0.86(0.08)	0.77(0.17)
ARFIMA38		0.96(0.18)	0.32(0.03)	1.03(0.22)	0.89(0.25)	0.84(0.10)	0.34(0.03)	0.85(0.08)	0.82(0.13)
ARFIMA39		1.00(0.17)	0.33(0.03)	1.04(0.21)	0.81(0.29)	0.84(0.11)	0.34(0.03)	0.86(0.08)	0.77(0.18)
ARFIMA40		1.06(0.18)	0.33(0.03)	1.06(0.21)	0.81(0.27)	0.87(0.10)	0.34(0.03)	0.89(0.07)	0.80(0.16)
ARFIMA41		1.04(0.19)	0.33(0.03)	1.06(0.22)	0.93(0.28)	0.87(0.11)	0.34(0.03)	0.89(0.08)	0.85(0.22)
ARFIMA42		1.05(0.19)	0.33(0.03)	1.04(0.22)	0.79(0.27)	0.88(0.10)	0.35(0.03)	0.90(0.08)	0.77(0.13)
ARFIMA43		1.32(0.17)	0.34(0.03)	1.34(0.14)	0.79(0.15)	1.20(0.13)	0.35(0.03)	1.26(0.10)	0.76(0.11)
ARFIMA44		1.30(0.18)	0.34(0.02)	1.34(0.14)	1.02(0.40)	1.21(0.12)	0.35(0.03)	1.26(0.09)	0.73(0.21)
ARFIMA45		1.32(0.17)	0.35(0.02)	1.35(0.14)	0.81(0.18)	1.22(0.12)	0.35(0.02)	1.27(0.09)	0.76(0.12)

$^{(*)}$ Notes: Models are estimated as discussed above, and model acronyms used are outlined in Table 1. d estimateds are based on GPH, RR, WHI and AML estimators. Table lists the average estimated value of d for ARFIMA DGPs over the 500 Monte Carlo iterations. Values in parenthesis are the standard deviation of the estimated d.

Table 7: Estimated d for Non-ARFIMA DGPs[(*)]

	Sample Size 500				Sample Size 2000			
Model	GPH	RR	WHI	AML	GPH	RR	WHI	AML
RW	1.18(0.16)	0.34(0.03)	1.24(0.18)	1.03(0.19)	1.16(0.10)	0.37(0.02)	1.22(0.11)	1.01(0.11)
AR1	-0.03(0.17)	0.06(0.04)	-0.04(0.21)	-0.05(0.23)	-0.02(0.11)	0.06(0.03)	-0.01(0.08)	-0.05(0.18)
AR2	-0.01(0.17)	0.10(0.04)	-0.03(0.21)	-0.12(0.23)	-0.01(0.10)	0.09(0.03)	0.00(0.08)	-0.03(0.12)
AR3	0.11(0.17)	0.17(0.04)	0.01(0.20)	-0.08(0.20)	0.04(0.10)	0.14(0.03)	0.05(0.07)	-0.02(0.12)
AR4	0.65(0.18)	0.27(0.04)	0.51(0.22)	0.67(0.37)	0.43(0.10)	0.24(0.03)	0.48(0.08)	0.28(0.46)
REG1	0.33(0.19)	0.19(0.04)	0.21(0.21)	0.10(0.47)	0.19(0.11)	0.17(0.03)	0.21(0.08)	0.19(0.26)
REG2	0.67(0.17)	0.25(0.06)	0.80(0.18)	0.74(0.35)	0.61(0.09)	0.26(0.03)	0.67(0.07)	0.63(0.27)
REG3	0.68(0.18)	0.25(0.09)	0.93(0.15)	0.77(0.37)	0.67(0.18)	0.28(0.11)	0.76(0.04)	0.82(0.34)
MA1	-0.04(0.17)	0.08(0.04)	-0.03(0.21)	-0.06(0.18)	-0.02(0.11)	0.07(0.03)	-0.01(0.08)	-0.02(0.09)
MA2	-0.04(0.18)	0.06(0.04)	-0.04(0.21)	-0.07(0.20)	-0.02(0.11)	0.05(0.03)	-0.01(0.08)	-0.02(0.10)
MA3	-0.04(0.18)	-0.02(0.04)	-0.05(0.20)	-0.10(0.20)	-0.02(0.11)	-0.02(0.03)	-0.02(0.08)	0.00(0.16)
MA4	-0.19(0.18)	-0.18(0.03)	-0.09(0.19)	-0.09(0.22)	-0.07(0.11)	-0.16(0.02)	-0.06(0.08)	-0.03(0.11)
GARCH1	-0.05(0.19)	0.02(0.04)	-0.05(0.21)	-0.05(0.18)	-0.02(0.11)	0.02(0.03)	-0.01(0.08)	-0.02(0.11)
GARCH2	-0.01(0.18)	0.11(0.04)	-0.03(0.21)	-0.11(0.24)	-0.01(0.11)	0.09(0.03)	0.00(0.08)	-0.02(0.13)
GARCH3	0.41(0.18)	0.23(0.04)	0.22(0.21)	0.22(0.38)	0.21(0.10)	0.21(0.03)	0.23(0.08)	-0.02(0.14)
ARX1	-0.04(0.18)	0.06(0.04)	-0.04(0.20)	-0.06(0.24)	-0.01(0.11)	0.06(0.03)	0.00(0.08)	-0.04(0.17)
ARX2	0.00(0.17)	0.11(0.04)	-0.03(0.20)	-0.11(0.23)	-0.01(0.11)	0.09(0.03)	0.00(0.08)	-0.02(0.12)
ARX3	0.12(0.17)	0.17(0.04)	0.03(0.22)	-0.07(0.21)	0.03(0.10)	0.14(0.03)	0.04(0.08)	-0.02(0.12)
ARX4	0.65(0.18)	0.27(0.04)	0.52(0.21)	0.67(0.36)	0.44(0.10)	0.24(0.03)	0.48(0.08)	0.25(0.46)
ARIMA1	-0.07(0.17)	-0.04(0.04)	-0.06(0.19)	-0.16(0.21)	-0.03(0.10)	-0.03(0.03)	-0.02(0.08)	-0.10(0.19)
ARIMA2	-0.03(0.17)	0.09(0.04)	-0.03(0.21)	-0.06(0.24)	-0.01(0.10)	0.08(0.03)	0.00(0.08)	-0.07(0.19)
ARIMA3	0.07(0.19)	0.11(0.04)	0.00(0.21)	-0.11(0.36)	0.02(0.11)	0.10(0.03)	0.03(0.08)	0.01(0.23)
ARIMA4	0.13(0.17)	0.18(0.04)	0.01(0.20)	-0.09(0.22)	0.03(0.10)	0.15(0.03)	0.04(0.08)	-0.04(0.12)
ARIMA5	1.26(0.16)	0.34(0.03)	1.35(0.15)	0.89(0.28)	1.17(0.09)	0.37(0.02)	1.23(0.12)	0.87(0.24)
ARIMA6	1.28(0.16)	0.34(0.02)	1.36(0.16)	1.04(0.26)	1.18(0.10)	0.37(0.02)	1.24(0.12)	0.90(0.22)
ARIMA7	1.32(0.15)	0.35(0.03)	1.36(0.14)	1.05(0.37)	1.20(0.10)	0.37(0.02)	1.26(0.12)	1.02(0.28)
ARIMA8	1.33(0.14)	0.35(0.03)	1.38(0.13)	0.93(0.27)	1.19(0.10)	0.37(0.02)	1.25(0.12)	1.00(0.17)

[(*)] Notes: See notes for Table 6.

Table 8: Relative MSFE Performance: ARFIMA DGPs[(*)]

| Model | d | Sample Size 500 | | Sample Size 2000 | |
		ARFIMA	Non-ARFIMA	ARFIMA	Non-ARFIMA
ARFIMA1	0.05	0.47 (0.32)	0.53 (0.06)	0.48 (0.26)	0.52 (0.07)
ARFIMA2		0.55 (0.35)	0.45 (0.11)	0.57 (0.38)	0.43 (0.11)
ARFIMA3		0.45 (0.15)	0.55 (0.07)	0.53 (0.11)	0.47 (0.11)
ARFIMA4		0.45 (0.33)	0.55 (0.05)	0.43 (0.23)	0.57 (0.08)
ARFIMA5		0.54 (0.33)	0.46 (0.10)	0.53 (0.12)	0.47 (0.10)
ARFIMA6		0.38 (0.14)	0.62 (0.06)	0.51 (0.13)	0.49 (0.12)
ARFIMA7		0.27 (0.23)	0.73 (0.09)	0.36 (0.20)	0.64 (0.09)
ARFIMA8		0.44 (0.15)	0.56 (0.08)	0.35 (0.11)	0.65 (0.06)
ARFIMA9		0.44 (0.14)	0.56 (0.16)	0.35 (0.15)	0.65 (0.05)
ARFIMA10	0.25	0.57 (0.17)	0.43 (0.07)	0.78 (0.19)	0.22 (0.05)
ARFIMA11		0.65 (0.18)	0.35 (0.08)	0.76 (0.17)	0.24 (0.02)
ARFIMA12		0.62 (0.02)	0.38 (0.09)	0.75 (0.20)	0.25 (0.01)
ARFIMA13		0.60 (0.07)	0.40 (0.06)	0.76 (0.23)	0.24 (0.02)
ARFIMA14		0.59 (0.12)	0.41 (0.07)	0.74 (0.17)	0.26 (0.02)
ARFIMA15		0.57 (0.02)	0.43 (0.09)	0.78 (0.20)	0.22 (0.01)
ARFIMA16		0.51 (0.03)	0.49 (0.03)	0.55 (0.09)	0.45 (0.04)
ARFIMA17		0.48 (0.17)	0.52 (0.03)	0.62 (0.08)	0.38 (0.06)
ARFIMA18		0.52 (0.05)	0.48 (0.05)	0.52 (0.08)	0.48 (0.05)
ARFIMA19	0.45	0.60 (0.09)	0.40 (0.04)	0.72 (0.20)	0.28 (0.01)
ARFIMA20		0.65 (0.14)	0.35 (0.07)	0.74 (0.20)	0.26 (0.02)
ARFIMA21		0.65 (0.22)	0.35 (0.01)	0.91 (0.39)	0.09 (0.00)
ARFIMA22		0.57 (0.08)	0.43 (0.05)	0.84 (0.31)	0.16 (0.00)
ARFIMA23		0.64 (0.10)	0.36 (0.13)	0.68 (0.24)	0.32 (0.02)
ARFIMA24		0.60 (0.11)	0.40 (0.03)	0.85 (0.34)	0.15 (0.01)
ARFIMA25		0.44 (0.03)	0.56 (0.05)	0.62 (0.09)	0.38 (0.03)
ARFIMA26		0.53 (0.03)	0.47 (0.10)	0.70 (0.10)	0.30 (0.02)
ARFIMA27		0.39 (0.03)	0.61 (0.08)	0.60 (0.12)	0.40 (0.03)
ARFIMA28	0.55	0.54 (0.08)	0.46 (0.05)	0.81 (0.31)	0.19 (0.01)
ARFIMA29		0.65 (0.14)	0.35 (0.08)	0.68 (0.19)	0.32 (0.03)
ARFIMA30		0.58 (0.13)	0.42 (0.04)	0.84 (0.32)	0.16 (0.01)
ARFIMA31		0.50 (0.09)	0.50 (0.07)	0.78 (0.29)	0.22 (0.01)
ARFIMA32		0.48 (0.06)	0.52 (0.24)	0.64 (0.18)	0.36 (0.12)
ARFIMA33		0.55 (0.12)	0.45 (0.05)	0.71 (0.20)	0.29 (0.02)
ARFIMA34		0.38 (0.02)	0.62 (0.07)	0.52 (0.06)	0.48 (0.05)
ARFIMA35		0.44 (0.03)	0.56 (0.06)	0.58 (0.07)	0.42 (0.04)
ARFIMA36		0.31 (0.02)	0.69 (0.08)	0.45 (0.06)	0.55 (0.07)
ARFIMA37	0.75	0.33 (0.05)	0.67 (0.12)	0.63 (0.15)	0.37 (0.04)
ARFIMA38		0.38 (0.03)	0.62 (0.23)	0.56 (0.11)	0.44 (0.08)
ARFIMA39		0.40 (0.07)	0.60 (0.11)	0.32 (0.06)	0.68 (0.17)
ARFIMA40		0.31 (0.03)	0.69 (0.11)	0.34 (0.07)	0.66 (0.11)
ARFIMA41		0.29 (0.01)	0.71 (0.33)	0.47 (0.04)	0.53 (0.14)
ARFIMA42		0.36 (0.03)	0.64 (0.13)	0.39 (0.07)	0.61 (0.12)
ARFIMA43		0.32 (0.02)	0.68 (0.07)	0.28 (0.03)	0.72 (0.10)
ARFIMA44		0.33 (0.02)	0.67 (0.08)	0.40 (0.03)	0.60 (0.05)
ARFIMA45		0.28 (0.01)	0.72 (0.11)	0.25 (0.01)	0.75 (0.13)

[(*)] Notes: Entries show the proportion of times different models are preferred based on comparison of point MSFE. Model acronyms used are outlined in Table 1. Values in the brackets show the same on the basis of Diebold and Mariano (DM) tests based on nominal 20% level critical values. Results are based on 500 Monte Carlo iterations.

Table 9: Relative MSFE Performance: Non-ARFIMA DGPs[(*)]

	Sample Size 500		Sample Size 2000	
Model	ARFIMA	Non-ARFIMA	ARFIMA	Non-ARFIMA
RW	0.22 (0.01)	0.78 (0.41)	0.15 (0.00)	0.85 (0.41)
AR1	0.44 (0.05)	0.56 (0.30)	0.42 (0.07)	0.58 (0.31)
AR2	0.41 (0.04)	0.59 (0.31)	0.49 (0.06)	0.51 (0.25)
AR3	0.41 (0.04)	0.59 (0.30)	0.36 (0.04)	0.64 (0.32)
AR4	0.24 (0.02)	0.76 (0.37)	0.18 (0.02)	0.82 (0.46)
MA1	0.54 (0.06)	0.46 (0.05)	0.50 (0.04)	0.50 (0.06)
MA2	0.47 (0.04)	0.53 (0.21)	0.50 (0.03)	0.50 (0.09)
MA3	0.49 (0.06)	0.51 (0.09)	0.52 (0.05)	0.48 (0.07)
MA4	0.48 (0.04)	0.52 (0.07)	0.50 (0.06)	0.50 (0.09)
GARCH1	0.36 (0.02)	0.64 (0.22)	0.37 (0.04)	0.63 (0.24)
GARCH2	0.14 (0.01)	0.86 (0.69)	0.11 (0.01)	0.89 (0.75)
GARCH3	0.12 (0.01)	0.88 (0.68)	0.09 (0.01)	0.91 (0.74)
ARX1	0.48 (0.05)	0.52 (0.26)	0.34 (0.19)	0.66 (0.32)
ARX2	0.44 (0.04)	0.56 (0.29)	0.42 (0.02)	0.58 (0.30)
ARX3	0.39 (0.03)	0.61 (0.32)	0.40 (0.03)	0.60 (0.31)
ARX4	0.26 (0.09)	0.74 (0.38)	0.24 (0.09)	0.76 (0.46)
ARIMA1	0.50 (0.08)	0.50 (0.06)	0.50 (0.04)	0.50 (0.06)
ARIMA2	0.47 (0.05)	0.53 (0.08)	0.48 (0.02)	0.52 (0.06)
ARIMA3	0.49 (0.04)	0.51 (0.27)	0.47 (0.03)	0.53 (0.21)
ARIMA4	0.43 (0.02)	0.57 (0.06)	0.42 (0.03)	0.58 (0.08)
ARIMA5	0.36 (0.02)	0.64 (0.17)	0.30 (0.01)	0.70 (0.17)
ARIMA6	0.33 (0.04)	0.67 (0.15)	0.30 (0.03)	0.70 (0.12)
ARIMA7	0.34 (0.03)	0.66 (0.21)	0.26 (0.02)	0.74 (0.18)
ARIMA8	0.25 (0.03)	0.75 (0.24)	0.20 (0.03)	0.80 (0.24)

[(*)] Notes: See notes for Table 8.