Applied Financial Economics, 2003, **13**, 495–502

Routledge
Taylor & Francis Group

Long memory and outliers in stock market returns

JUSSI TOLVI

Department of Economics, 20014 University of Turku, Finland
E-mail: jussi.tolvi@utu.fi

Long memory in the form of fractional integration is analysed in stock market returns. Special emphasis is placed on taking into account the potential bias caused by neglected outliers in the data. It is first shown by a simulation experiment that outliers will bias the estimated fractional integration parameter towards zero. In a monthly data set, consisting of stock market indices of 16 OECD countries, statistically significant long memory is found for three countries. In one of these long memory is only found when outliers are first taken into account.

I. INTRODUCTION

Long memory in time series data manifests itself as temporal dependence over long periods of time. Usually this is in economics considered as significant autocorrelations at long lags, of up to hundreds of periods. Such models have recently been examined quite extensively in the theoretical literature (see Baillie, 1996, for an introduction). However, one aspect mostly missing from this work is the effect of outliers on these models. It is by now well known that outliers will bias estimators in several other time series models, and there is no reason to assume that the same would not be true also of long memory models. The long memory model that will be considered in this study is the autoregressive fractional integration moving average (ARFIMA) model, where the fractional integration parameter determines the long memory properties of the data.

The closest analogy with estimating the fractional integration parameter is perhaps that of estimating an autoregressive parameter. In that case it has been clearly established that the presence of certain types of outliers will bias the traditional (nonrobust) estimates towards zero (see, e.g., Tsay, 1986; Chen and Liu, 1993). Intuitively it seems likely that this could happen also in fractional integration models, although proving it analytically may not be easy. In this study therefore, this issue is first examined by some Monte Carlo simulation experiments.

There have been several applications of long memory models to empirical economic data, including stock market returns. Earlier research has mostly used a semiparametric estimator by Geweke and Porter-Hudak (1983), which however suffers from a drawback. The application of this estimator requires a choice for the number of periodogram ordinates to be used in the estimation, but it is not clear which value should be used in practice. Different values usually give somewhat different estimates. In addition, in the presence of short range dependencies (say, AR or MA terms in the data generating process) the GPH estimator is known to be biased in small samples (Agiakloglou *et al.*, 1992). Baillie (1996) provides a further discussion on these topics. In this study more recently proposed maximum likelihood estimation methods will be used. These have better small sample properties, and will be discussed briefly in the next section.

In the empirical part of this paper monthly stock market indices from 16 OECD countries, and a daily US stock market index will be examined by estimating ARFIMA models for them. Earlier research, in trying to find statistically significant long memory for this kind of data, has come up with mixed results. Nevertheless, at least some evidence of long memory has been found for some monthly, weekly and daily stock market indices by Crato (1994), Cheung and Lai (1995), Barkoulas and Baum (1996), Barkoulas *et al.* (2000) and Sadique and Silvapulle (2001).

Smaller markets in the monthly data set may be the most likely place to find significant long memory. Since small markets do not always seem to behave as expected, and may be less efficient than larger markets, it is possible

Applied Financial Economics ISSN 0960–3107 print/ISSN 1466–4305 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/09603100210161983

that long memory will be detected in them. This point was raised by Barkoulas *et al.* (2000), who examined weekly returns in the Greek stock market during the 1980s, and found clear evidence of statistically significant long memory.

Furthermore, if the presence of outliers biases the results for long memory tests and estimates, as will be shown later to be the case for ARFIMA estimates, it is also possible that significant long memory is hidden by outliers, and has therefore not been reliably detected so far. To see whether this is indeed the case, models where potential outliers have been taken into account with dummy variables will also be estimated for the same data. By comparing the two sets of results (i.e. from the basic and the outlier models), it can be seen whether the outliers make a difference with regard to inference about long memory.

II. THE ARFIMA MODEL AND OUTLIERS

The long memory model used in this paper is the standard autoregressive fractional integration moving average, or ARFIMA(p, d, q), model. For the observed series y_t it is given by

$$\Phi_p(L)(1-L)^d(y_t-\mu) = \Theta_q(L)\varepsilon_t, \quad t=1,\ldots,T \tag{1}$$

where L is the lag operator $(L^j y_t = y_{t-j})$, $\Phi_p(L) = 1-\phi_1 L - \ldots - \phi_p L^p$ is the autoregressive, and $\Theta_q(L) = 1+\theta_1 L + \ldots + \theta_q L^q$ the moving average lag polynomial. The differencing parameter d need not be an integer, but integer values of d lead to traditional ARIMA models. The fractional differencing operator $(1-L)^d$ is defined for non-integer d by a binomial expansion

$$(1-L)^d = \sum_{j=0}^{\infty}\binom{d}{j}(-L)^j \tag{2}$$

In addition, the usual assumptions that $\varepsilon_t \sim NID(0, \sigma^2)$, that all roots of the AR and MA polynomials are outside the unit circle, and that they do not have common roots, will be made.

The long range properties of such series depend on the value of d. For $d \in (0, 0.5)$ the (theoretical) autocorrelations are all positive. They decay hyperbolically to zero as the lag length increases, compared to the usual exponential decay in the case of a stationary ARMA model with $d = 0$. This property is also one of the definitions of long memory. For $d \in (-0.5, 0)$ the series is said to exhibit intermediate memory. In this case the autocorrelations are all negative, and decay hyperbolically to zero. For $d \geqslant 0.5$ the series are no longer covariance stationary. For a more detailed discussion see, for example, Baillie (1996) or Ooms and Doornik (1999).

It will be assumed that the observed data consists of an underlying ARFIMA series, and possibly a few outliers that occur due to exogenous shocks at random points in time. The study will distinguish between four outlier types, described below. Such outliers can be taken into account in estimation by the use of appropriate dummy variables. In practical work, the potential dummy variables have of course first to be somehow identified. In this study the series will first be approximated with ARMA models, and an outlier detection procedure used to detect all significant outliers in the fitted ARMA models. The detected outliers will then be incorporated into the ARFIMA models with corresponding dummies, and their statistical significance tested in the final ARFIMA models.

The ARFIMA model with dummy variables, called the outlier model from now on, can be written as

$$\Phi_p(L)(1-L)^d(y_t - x_t'\beta) = \Theta_q(L)(\varepsilon_t + z_t'\gamma) \tag{3}$$

where dummies for two different kinds of outlier types are included. First, the X matrix includes dummy variable vectors for additive outlier types. These can be thought to occur as if on top of the underlying series. The additive outlier types are a single additive outlier (AO) which only affects one observation, a permanent level shift (LS) which affects all following observations, and a temporary change (TC) outlier, whose effect lies between these two extremes. As is usual, temporary change outliers are here assumed to die out exponentially, and the values of the dummy variable from the date of the outlier on are $1, 0.7, 0.7^2, 0.7^3, \ldots$. Dummy variable vectors for the second kind of outlier, so called innovational outliers (IOs), are included in the Z matrix. This outlier type can be thought of as an abnormal shock (in the ε_t variable), the effect of which is propagated, via the data generating ARFIMA process, into the following observations as well. See Chen and Liu (1993) for a more thorough discussion on these outlier types.

The estimation of ARFIMA models will be done with the Arfima package version 1.0 (Doornik and Ooms, 1999) for Ox (Doornik, 1998).[1] The Arfima package allows the use of three parametric estimation methods. These are the exact maximum likelihood (EML), nonlinear least squares (NLS) and modified profile maximum likelihood (MPL). See Doornik and Ooms (1999) for the details of these methods and further references on them. In addition, Hauser (1999) examines the small sample properties of the EML and MPL estimators in various situations. The outlier robustness of these estimators is examined in the next section.

It seems that for simple ARFIMA(1, d, 1) models without outliers, the MPL method is preferable to the EML

[1] See also the companion article by Ooms and Doornik (1999), which discusses the implementation of the software.

method (Ooms and Doornik, 1999). Unfortunately, the MPL method in the current version of Arfima package for Ox does not allow dummies for innovational outliers in the model, and therefore the EML estimation method will be used in the empirical part of this paper.

III. OUTLIER ROBUSTNESS OF LONG MEMORY ESTIMATES

It is already known that level shifts cause problems for accurate estimation of ARFIMA model parameters (Bos *et al.*, 1999). For other outlier types similar studies have apparently not yet been done, apart from Beran (1994), who considers the effects of additive outliers on the exact maximum likelihood estimator, and proposes a robust estimator as an alternative. However, his results are based on a very limited, and, with respect to the magnitudes of the outliers, somewhat unlikely simulation design. In this section the effects of additive outliers are therefore examined with some more extensive simulation experiments.

The simulation results presented here are for samples of 500 observations, which is a moderate sample size for analysing long memory. The true fractional differencing parameter $d = -0.4, -0.3, \ldots, 0.3, 0.4$, the mean of the series is set to zero, and the error variance σ^2 equal to one. To generate fractionally integrated series, the Choleski decomposition method is used (see Doornik and Ooms, 1999). Outliers, both negative and positive with equal probabilities, are then added to the series so that each observation is an outlier with probabilities 0 (i.e. no outliers), 0.5, 1, 2.5 and 5%, and the outlier magnitudes used are 3, 5 and 7. As will be seen later, these outlier probabilities are reasonable in the sense that the amounts of detected outliers in the empirical data are quite similar to the ones used in the simulations. The outlier magnitudes, on the other hand, may even be viewed as somewhat conservative, compared to the largest observations in the empirical data, for example during the stock market crash of 1987.

The data generating models have no autoregressive or moving average terms, and are therefore all ARFIMA$(0, d, 0)$, or FI(d) models. It would obviously be interesting to examine the effects of outliers also on models where there are autoregressive and moving average terms, but I will at the moment concentrate only on the simplest situation. Note also, that in an ARFIMA$(0, d, 0)$ model, there is no difference between additive and innovational outliers. The effects of temporary changes and level shifts will also not be considered here.

In the estimation stage, the true model is assumed to be known, along with the true mean of the series. This is of course an ideal situation, and unlikely in practice. However, it is adequate for our purposes, since the interest here is only on the effects of outliers. All simulations are based on 1000 replications. It should perhaps also be noted, that the estimation algorithms failed to converge in only a handful of the simulated samples. For this sample size they are therefore very reliable.

Table 1 has the mean of all of the replications for the EML estimator.[2] As can be seen from the table, in the absence of outliers the EML estimator is unbiased. When outliers occur in the data, however, the estimator becomes

Table 1. *Simulated mean estimates of the long memory parameter*

		d								
λ	ω	−0.4	−0.3	−0.2	−0.1	0.0	0.1	0.2	0.3	0.4
0	0	−0.403	−0.304	−0.202	−0.102	−0.001	0.099	0.197	0.295	0.398
0.005	3	−0.353	−0.275	−0.190	−0.097	−0.003	0.092	0.192	0.287	0.387
	5	−0.306	−0.240	−0.170	−0.090	−0.004	0.088	0.181	0.278	0.373
	7	−0.261	−0.213	−0.151	−0.080	−0.003	0.082	0.169	0.257	0.352
0.01	3	−0.319	−0.255	−0.176	−0.094	−0.003	0.090	0.187	0.282	0.377
	5	−0.253	−0.207	−0.149	−0.080	0.000	0.083	0.169	0.258	0.353
	7	−0.201	−0.164	−0.120	−0.067	−0.003	0.069	0.148	0.233	0.323
0.025	3	−0.258	−0.211	−0.152	−0.081	−0.003	0.082	0.172	0.262	0.354
	5	−0.178	−0.147	−0.107	−0.059	−0.002	0.064	0.141	0.220	0.305
	7	−0.125	−0.102	−0.078	−0.043	−0.003	0.047	0.109	0.180	0.264
0.05	3	−0.203	−0.166	−0.118	−0.066	−0.002	0.070	0.149	0.233	0.326
	5	−0.122	−0.102	−0.074	−0.043	−0.002	0.048	0.110	0.180	0.260
	7	−0.076	−0.064	−0.048	−0.028	−0.003	0.032	0.077	0.134	0.211

Note: The results are based on 1000 replications of exact maximum likelihood estimation. The sample size is 500, d is the true fractional integration parameter, λ the probability of an outlier occurring at each observation, and ω the outlier magnitude.

[2] The NLS and MPL estimators give essentially the same results, and the results for these two estimators will therefore not be tabulated here to save space. They are, however, as well as other unreported results, available from the author.

biased towards zero. More and larger outliers cause also larger biases. This is a clear trend and can be seen with all parameter combinations. The standard deviations of the estimates also grow in the presence of outliers. Note also, that the biases are larger for negative values of d.

The same experiments were also repeated for samples of 100 observations. These results will, however, not be reported here. Suffice it to note, that the biases of the estimators are roughly similar to those with 500 observations, especially for positive values of d. In addition, with less observations there is more variability in the results, and the standard deviations of the estimates are also larger. Overall, however, roughly the same conclusions can be drawn from both sample sizes.

Although the outlier induced biases are not as dramatic as for some other estimation situations, they are nevertheless notable. With larger and more frequent outliers the estimates approach zero, which was also the finding of Beran (1994). It seems therefore warranted to conclude that it is essential to take any potential outliers into account when estimating ARFIMA models – this is of course true with any other model as well. Sometimes this is done in the literature at least informally (in, e.g. Ooms and Doornik, 1999), but mostly, and unfortunately, it seems not.

IV. EMPIRICAL RESULTS

The data series in this section are monthly differenced logarithms of stock price indices for 16 OECD countries, mostly from 1960 to 1999, which gives roughly 475 observations. Full details of the series, and the data source, are given in the Appendix.

Potential outliers were first identified from each series with the program TRAMO (Gómez & Maravall, 1994, 1995), which uses a slight modification of the Chen and Liu (1993) outlier detection algorithm. If one can assume that a time series can be approximated with an ARMA model, this procedure is commonly used to detect any abnormal observations. In this study the use of such a black box method can be motivated by the fact that the main interest is on examining the robustness (with respect to a few isolated outliers) of the inferences regarding the long memory properties of stock returns. The outlier detection algorithm will find the most unusual observations, which can then be taken into account in the final ARFIMA models.

In addition to outlier detection, the automatic model selection option of TRAMO, which is based on the Schwarz information criteria, was also used. The results gave therefore also preliminary AR and MA lag lengths to use in the next stage of estimating the ARFIMA models. The critical value used in the outlier detection was 4.0. This is considered a low sensitivity value for this sample size, meaning that only the most outlying observations will be detected as outliers. All four outlier types (AO, LS, TC, IO) were searched for.

The dates, types and signs of the detected outliers are given in Table 2. Note first that no level shifts were detected in any of the series. The number of outliers per series varies from one (Germany and Japan) to eleven (Ireland), the

Table 2. *Detected outliers*

	Outlier dates	Outlier types
Australia	68:5, 74:6, 87:10, 87:11	$IO+, TC-, IO-, AO-$
Germany	87:11	$AO-$
Belgium	85:4, 87:11	$AO-, IO-$
Canada	80:3, 87:10, 98:8	$IO-, AO-, IO-$
Denmark	72:10, 73:11, 80:10,	$TC+, AO-, IO+,$
	83:1, 83:2, 83:8	$AO-, TC+, IO+$
Spain	86:3, 87:11	$TC+, IO-$
Finland	66:9, 68:4, 68:6, 87:1,	$AO-, IO+, AO+, AO-,$
	92:10, 98:8, 98:11	$TC+, TC-, TC+$
France	82:3, 82:4, 87:10, 88:2	$AO-, AO+, IO-, AO+$
Ireland	74:10, 75:2, 75:7, 77:1,	$IO-, TC+, AO-, TC+,$
	86:3, 87:11, 87:12, 88:2,	$AO+, IO-, AO-, IO+,$
	90:9, 91:3, 98:9	$IO-, AO+, AO-$
Italy	81:7, 94:4, 94:5, 94:7	$AO-, AO-, TC+, TC-$
Japan	92:9	$AO+$
The Netherlands	75:1, 81:9, 83:1, 87:11	$IO+, AO-, AO-, IO-$
Norway	74:9, 83:1, 87:11	$IO-, TC+, IO-$
Sweden	87:10, 90:9, 92:11	$IO-, IO-, AO+$
UK	73:12, 74:6, 75:1,	$AO-, TC-, IO+,$
	75:2, 87:11	$IO+, AO-$
USA	62:6, 87:10	$AO-, TC-$

Note: The outlier type indicates also the sign of each outlier $(+/-)$.

mean being just under four. The share of outliers (out of all observations) varies therefore from 0.2% to 2.2%, which is not an unreasonable amount of outlying observations in financial data, and agrees also with the outlier probabilities used in the simulation experiment of the previous section. Roughly two-thirds of the detected outliers are negative, which is an indication of negative skewness in the series. And as could be expected, October/November 1987 is detected as an (negative) outlier in most series. These detected outliers were then used in the next stage as dummies in the estimated ARFIMA outlier models (dummies for AOs and TCs as X variables, and for IOs as Z variables; see Equation 3).

The ARMA model selected for the series was usually MA(1), apart from the following exceptions. An ARMA(0,0) model was selected for Canada, Finland, Italy and Norway, and an AR(1) model for Denmark and Ireland. Full estimation results will not be given here to save space, since the main interest is on the fractional integration parameters. Suffice it to note, that the dummy variables were always statistically highly significant, usually at levels below 0.1%, and the ARMA parameters were significant at least at the 10% level. The outlier magnitudes, or the estimated parameter values of the dummies, were larger than 0.1 in absolute value, and many were over 0.2. For comparison, the standard deviations of the series range mostly from 0.04 to 0.05. Observations as large as seven times the series' standard deviations are therefore not unusual in data of this kind. The residual diagnostics, namely normality, ARCH and remaining autocorrelation, were quite often significant. However, often the hypothesis of normality would not have been rejected in the outlier model, which is of course due to the normalizing effect of removing the outliers.

Table 3 gives the estimation results. In addition to the selected ARMA lag orders and fractional integration parameter estimates, the residual standard deviations are given for the basic and the outlier models. The last column gives also the likelihood ratio (LR) for the two models. Examining first the results for the basic models without

Table 3. *Estimated ARFIMA models*

		Basic model		Outlier model		
	ARMA	$\hat{d}_B$	$\hat{\sigma}_B$	$\hat{d}_O$	$\hat{\sigma}_O$	LR
Australia	(0,1)	−0.073 (0.053)	0.051	−0.0086 (0.057)	0.046	103.94
Germany	(0,1)	0.048 (0.052)	0.038	0.043 (0.050)	0.036	48.062
Belgium	(0,1)	−0.051 (0.047)	0.042	−0.014 (0.044)	0.039	73.070
Canada	(0,0)	−0.056 (0.039)	0.049	−0.039 (0.036)	0.045	72.004
Denmark	(1,0)	0.13 (0.058)	0.050	0.10 (0.056)	0.045	106.18
Spain	(0,1)	0.019 (0.048)	0.051	−0.0028 (0.049)	0.049	45.931
Finland	(0,0)	0.15 (0.038)	0.053	0.17 (0.035)	0.045	153.81
France	(0,1)	−0.054 (0.053)	0.056	−0.045 (0.048)	0.050	100.94
Ireland	(1,0)/(2,0)	−0.064 (0.098)	0.053	0.31 (0.069)	0.042	217.50
Italy	(0,0)	−0.025 (0.038)	0.091	0.063 (0.038)	0.070	223.20
Japan	(0,1)	0.053 (0.049)	0.040	0.066 (0.049)	0.040	16.322
The Netherlands	(0,1)	−0.071 (0.051)	0.055	0.023 (0.049)	0.040	292.19
Norway	(0,0)	0.041 (0.037)	0.066	0.013 (0.035)	0.060	95.886
Sweden	(0,1)	−0.0040 (0.053)	0.051	0.044 (0.054)	0.048	69.315
UK	(0,1)	−0.053 (0.048)	0.043	0.013 (0.047)	0.037	142.86
USA	(0,1)	−0.047 (0.057)	0.042	−0.012 (0.057)	0.041	26.755

Note: ARMA denotes the lag orders of the estimated models (basic model/outlier model in the case of Ireland), and LR is the logarithmic likelihood ratio for the two models.

the dummies, given in the second and third column of Table 3, statistically significant (at the 5% level) estimates of d are obtained only for Denmark and Finland. For the other series no clear indication of long memory can be found.

Results for the outlier models are given in the fourth and fifth columns of Table 3. The estimated ARMA parameters did not usually change much when the outliers were taken into account with dummies. The same is true for the fractional integration parameter. For Finland and Denmark, the countries with significant long memory in the basic model, the estimates of d increase and decrease slightly, respectively, but remain statistically significant. Most of the residual standard deviations, on the other hand, are considerably decreased when the dummies are added to the models. The dummies therefore explain a considerable amount of the variation in the data, which is also demonstrated by the clearly statistically significant LR test statistics.

All in all there are therefore no major changes in the estimates of d, and in the conclusions that can be drawn from them. The one notable exception is Ireland, for which there is overwhelming evidence of statistically significant long memory in the outlier model. The estimated d grows from -0.064 to 0.31 when the dummies are added to the model. In this case, therefore, taking the outliers into account changes the inference regarding long memory dramatically. Also, in this case, different AR lag orders are selected in the two ARFIMA models. The whole model is therefore changed when the outliers are taken into account. The influence of the outliers in the Irish data can also be seen from Fig. 1, which plots first the original return series, and then residuals, both from the basic and the outlier ARFIMA models.[3,4]

V. CONCLUSIONS

The simulation experiment in this paper shows clearly that neglected outliers may result in biased estimates of the fractional integration parameter. There seems to be no difference in the bias between the most popular estimation methods. Furthermore, the empirical results for stock market returns show that this possibility is not only theoretical. The influence of outliers does not seem to be great in most cases, but for the Irish monthly data they were decisive for inference regarding the fractional integration parameter. And as hypothesized, in the monthly data long memory is more likely found in smaller stock markets than in larger ones. Statistically significant fractional integration parameters were indeed found only for Denmark, Finland and Ireland, which are all small markets.

Based on these results, it seems therefore imperative that any potential outliers are taken into account also in estimating ARFIMA models. Nevertheless, the ultimate test of whether taking outliers into account in an analysis such as this one has any value, depends naturally on the final use of the model. If, for example, the aim is in the forecasting of stock market returns, it matters little whether some parameter estimate is statistically significant or not at some arbitrary level. To justify using an outlier model for this purpose, one should first demonstrate that the forecasts are improved if an outlier model is used instead of the basic one. In this study, this issue has not been tackled at all, but earlier research suggests that there may be at least some value in using ARFIMA models in forecasting over longer horizons (Barkoulas *et al.*, 2000).

Further research should perhaps consider other kinds of returns as well, for example for individual stocks, and for different data frequencies. Since aggregation, both temporal and cross-sectional, tends to decrease the amount and magnitude of outliers in the data, moving to individual stocks could produce interesting results. Another interesting topic is the possible long memory in absolute values of the returns (as a measure of volatility). The present availability of easy-to-use computer software has made this kind of work relatively easy. As for theoretical work, further research on the effects of outliers for ARFIMA models seems also warranted.

[3] One improvement to this analysis would be to include a GARCH specification for the errors of these ARFIMA models, since the variance of the series is clearly not constant. This could also have an effect on the outlier detection, so that less outliers could be detected in ARFIMA-GARCH models than what was found earlier. To consider this possibility, an attempt was made to estimate ARFIMA(p, d, q)-GARCH(1,1) models for the three series with statistically significant long memory. The models were estimated first without any dummies, and then with dummies for the outliers detected above, the idea here being that any statistically insignificant dummies can then be dropped from the model. In some of the models the estimation procedure failed to converge, even after some experimentation with different starting values and optimization parameters. This is no doubt due to the increased complexity of the models, relative to the rather small number of observations. All in all it seemed, however, that taking the GARCH structure in the series into account had very little effect on the point estimates of the long memory parameter and on the significance of the dummies.

[4] The same analysis was also carried out for daily returns of the Dow Jones Industrial Average index. Data from 1986 to 1999 was first divided to seven periods of 500 observations each. Full results for these data will not be presented in the text, but some brief comments on the findings can be made. The number of detected outliers in these samples varied from one to eight. Statistically significant values of d were found in the basic model for three samples: these were all negative, indicating intermediate memory. In the outlier models, nearly significant estimates were found in two additional samples. Taking the outliers into account therefore had some influence on the long memory estimation results in this data as well.

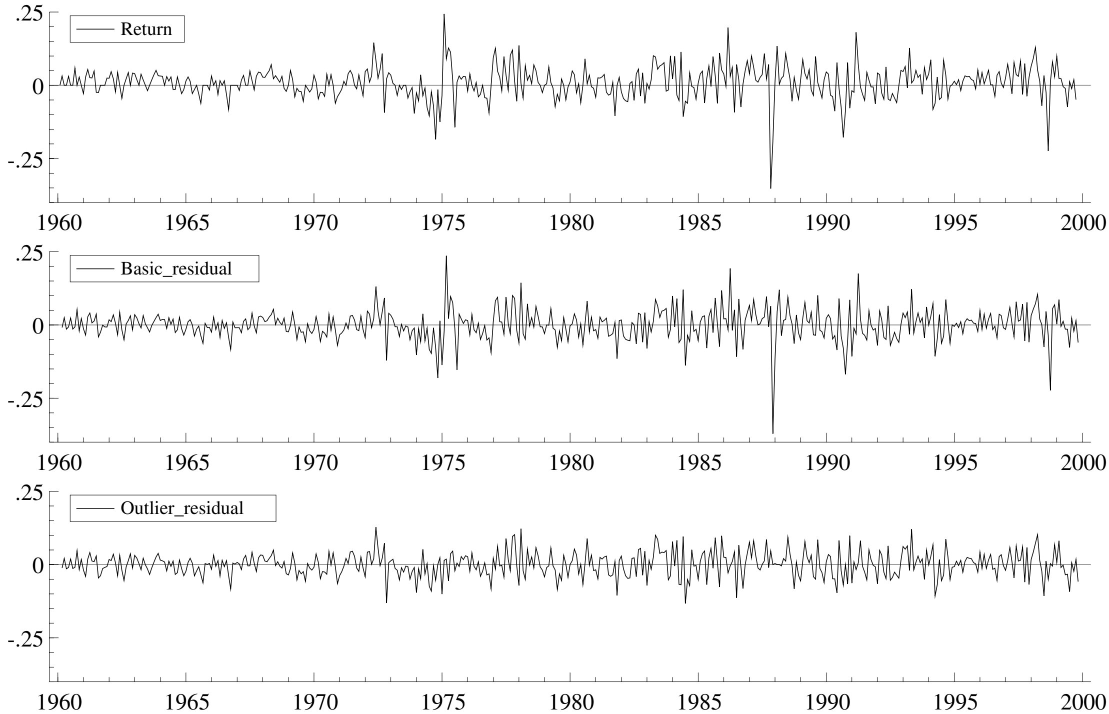

Fig. 1. *Returns and ARFIMA residuals (basic and outlier models) for Ireland*

ACKNOWLEDGEMENTS

The author would like to thank Markku Lanne, Marius Ooms and Seppo Pynnönen for comments on earlier versions of this paper.

REFERENCES

Agiakloglou, C., Newbold, P. and Wohar, M. (1992) Bias in an estimator of the fractional difference parameter, *Journal of Time Series Analysis*, **14**, 235–46.

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

Barkoulas, J. T. and Baum, C. F. (1996) Long term dependence in stock returns, *Economics Letters*, **53**, 253–9.

Barkoulas, J. T., Baum, C. F. and Travlos, N. (2000) Long memory in the Greek stock market, *Applied Financial Economics*, **10**, 177–84.

Beran, J. (1994) On a class of M-estimators for Gaussian long-memory models, *Biometrika*, **81**, 755–66.

Bos, C. S., Franses, P. H. and Ooms, M. (1999) Long memory and level shifts: Re-analyzing inflation rates, *Empirical Economics*, **24**, 427–49.

Chen, C. and Liu, L.-M. (1993) Joint estimation of model parameters and outlier effects in time series, *Journal of the American Statistical Association*, **88**, 284–97.

Cheung, Y.-W. and Lai, K. S. (1995) A search for long memory in international stock market returns, *Journal of International Money and Finance*, **14**, 597–615.

Crato, N. (1994) Some international evidence regarding the stochastic behavior of stock returns, *Applied Financial Economics*, **4**, 33–9.

Doornik, J. A. (1998) *Object-oriented matrix programming using Ox 2.0*, Timberlake Consultants, London.

Doornik, J. A. and Ooms, M. (1999) *A package for estimating, forecasting and simulating arfima models: Arfima package 1.0 for Ox* (An unpublished manuscript). Nuffield College, Oxford and Erasmus University Rotterdam. (See http://www.nuff.ox.ac.uk/Users/Doornik/)

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long memory time series models, *Journal of Time Series Analysis*, **4**, 221–38.

Gómez, V. and Maravall, A. (1994) *Program TRAMO. 'Time series regression with ARIMA noise, missing observations, and outliers.' Instructions for the user* (EUI working paper No. ECO 94/31), European University Institute, Florence.

Gómez, V. and Maravall, A. (1995) *Programs TRAMO and SEATS. Update: December 1995* (EUI working paper No. ECO 95/46), European University Institute, Florence.

Hauser, M. A. (1999) Maximum likelihood estimators for ARMA and ARFIMA models: A Monte Carlo study, *Journal of Statistical Planning and Inference*, **80**, 229–55.

Ooms, M. and Doornik, J. A. (1999) *Inference and forecasting for fractional autoregressive integrated moving average models, with an application to US and UK inflation* (Research Report No. 9947/A), Econometric Institute, Erasmus University Rotterdam.

Sadique, S. and Silvapulle, P. (2001) Long-term memory in stock market returns: international evidence, *International Journal of Finance and Economics*, **6**, 59–67.

Tsay, R. S. (1986) Time series model specification in the presence of outliers, *Journal of the American Statistical Association*, **81**, 132–41.

APPENDIX: DATA SOURCE AND SAMPLES

The monthly data is from the OECD database. Series codes and the samples are as follows. Australia, series AUOCSPRC (ASE all ordinaries), from 1960:1 to 1999:9. (West) Germany, series BDOCSPRC (CDAX) from 1960:1 to 1999:10. Belgium, series BGOCSPRC (BSE all shares), 1960:1 to 1999:8. Canada, series CNOCSPRC (TSE 300 composite), from 1960:1 to 1999:9. Denmark, series DKOCSPRC (CSE all shares) from 1960:1 to 1999:9. Spain, series ESOCSPRC (MSE general index) from 1961:1 to 1999:10. Finland, series FNOCSPRC (HEX all shares) from 1960:1 to 1999:9. France, series FROCSPRC (SBF 250), from 1960:1 to 1999:10. Ireland, series IROCSPRC (ISEQ index overall), from 1960:1 to 1999:10. Italy, series ITOCSPRC (Milan stock exchange) from 1960:1 to 1995:2. Japan, series JPOCSPRC (TSE TOPIX) from 1960:1 to 1999:9. The Netherlands, series NLOCSPRC (CBS all shares) from 1960:1 to 1999:8. Norway, series NWOCSPRC (Oslo stock exchange) from 1960:1 to 1998:5. Sweden, series SDOCSPRC (AFGX index) from 1960:1 to 1999:8. United Kingdom, series UKOCSPRC (FTSE actuaries non-financial) from 1960:1 to 1999:9. USA, series USOCSPRC (S&P industrials) from 1960:1 to 1993:6.