Forecasting irish inflation using ARIMA models

Meyler, Aidan and Kenny, Geoff and Quinn, Terry

Central Bank and Financial Services Authority of Ireland

December 1998

TECHNICAL PAPER

3/RT/98 DECEMBER 1998

FORECASTING IRISH INFLATION USING ARIMA MODELS

BY

AIDAN MEYLER*, GEOFF KENNY AND TERRY QUINN

The views expressed in this paper are not necessarily held by the Central Bank of Ireland and are the personal responsibility of the authors. The authors would like to thank John Frain and participants at an internal seminar in the Central Bank for comments on an earlier draft of this paper. All remaining errors and omissions are the authors. Comments and criticisms are welcome.

Economic Analysis, Research and Publications Department, Central Bank of Ireland, PO Box 559, Dublin 2. *Author for correspondence: Aidan Meyler. Email: enquiries@centralbank.ie

ABSTRACT

This paper outlines the practical steps which need to be undertaken to use autoregressive integrated moving average (ARIMA) time series models for forecasting Irish inflation. A framework for ARIMA forecasting is drawn up. It considers two alternative approaches to the issue of identifying ARIMA models - the Box Jenkins approach and the objective penalty function methods. The emphasis is on forecast performance which suggests more focus on minimising out-of-sample forecast errors than on maximising in-sample ‘goodness of fit’. Thus, the approach followed is unashamedly one of ‘model mining’ with the aim of optimising forecast performance. Practical issues in ARIMA time series forecasting are illustrated with reference to the harmonised index of consumer prices (HICP) and some of its major sub-components.

1. INTRODUCTION

The primary focus of monetary policy, both in Ireland and elsewhere, has traditionally been the maintenance of a low and stable rate of aggregate price inflation as defined by commonly accepted measures such as the consumer price index. The underlying justification for this objective is the widespread consensus, supported by numerous economic studies[1], that inflation is costly insofar as it undermines real, wealth-enhancing, economic activity.

From the beginning of 1999, the Irish economy faces a new environment in which monetary policy will be set by the Governing Council of the European Central Bank (ECB). The ECB is committed to a monetary policy which has the primary objective of maintaining price stability throughout the eleven euro-area countries **as a whole**.[2] Regardless of the exact strategy adopted by the ECB in the formulation of monetary policy, i.e., targeting monetary aggregates such as the broad money stock or direct inflation targeting, the provision of optimal and timely inflation forecasts represents a key ingredient in designing monetary policies which are geared toward the achievement of price stability. While it could be argued that Ireland's weight in the overall euro-area price index is relatively small and, as such, Irish inflation no longer warrants rigorous examination, it is important to note that Ireland has an input into monetary policy decision making at the ECB that is disproportionate to its economic size.

However, a more compelling argument for a continued focus on forecasting Irish inflation is the increased importance of fiscal policy and wage bargaining negotiations in the absence of independent monetary control. Arguably, the inflation forecast should be given greater weight in fiscal policy and in wage negotiations in Ireland than has been the case heretofore. Furthermore, given the possibility of sustained

---

[1] Two recent examples are Feldstein (1996) and Dotsey and Ireland (1996). Both of these studies argue, in particular, that even low rates of inflation of the order of 2% to 4% are highly costly over the long-run.

[2] "... the Governing Council of the ECB makes it clear that it will base its decisions on monetary, economic and financial developments in the euro area as a whole. The single monetary policy will adopt a euro area-wide perspective; it will not react to specific regional or national developments" (ECB, 1998).

differences in inflation rates across euro area currencies and the subsequent impact on competitiveness, monitoring and understanding price developments in individual economies will remain of significant importance.

In past studies of inflation by the Central Bank of Ireland the emphasis has been on testing economic theory and on empirical analysis. Even though some of these studies have been used as an input into the forecasting process within the Bank, they have not heretofore been subject to rigorous forecast evaluation techniques. This paper and Kenny *et al* (1998) set out to redress this deficiency and explicitly use time series techniques solely for forecasting purposes.

There are a number of approaches available for forecasting economic time series. One approach, which includes only the time series being forecast, is known as univariate forecasting. Autoregressive integrated moving average (ARIMA) modelling is a specific subset of univariate modelling, in which a time series is expressed in terms of past values of itself (the autoregressive component) plus current and lagged values of a 'white noise' error term (the moving average component). This paper focuses on ARIMA models. An alternative approach is multivariate time series forecasting. Multivariate models may consist of single equation models with exogenous explanatory variables or alternatively may include a structural or non-structural system of equations. Parallel research is also being currently undertaken within the Central Bank of Ireland into the use of Bayesian Vector Autoregressive (BVAR) models for forecasting Irish inflation (see Kenny *et al*, *op. cit.*).

In practice the formal econometric models outlined above are often supplemented by subjective 'off-model' inputs. Such information may include survey data gathered from liaising with retailers and manufacturing enterprises. Thus, inflation forecasting is an art rather than a hard science combining formal econometric techniques with forecasters' experience and expertise.

Practical issues in relation to ARIMA time series forecasting are illustrated using the harmonised index of consumer prices (HICP) and some of its major sub-components. The HICP was developed to allow comparison of inflation rates across EU states.

Prior to the development of the HICP, each state constructed its own consumer price index (CPI), which could differ in how they treated certain items such as housing, health, education and insurance. Previous work examining Irish inflation has concentrated on the CPI. This paper offers an opportunity to apply univariate techniques to the Irish HICP for the first time.[3]

In this paper six time series are examined. These allow us to consider many of the issues that arise in ARIMA time series forecasting. These series are the **overall Harmonised Index of Consumer Prices** (HICP) and the HICP broken down between **unprocessed foods** (HICPA), **processed food** (HICPB), **non-energy industrial goods** (HICPC), **energy** (HICPD), and **services** (HICPE). The emphasis is on forecast performance which suggests more focus on minimising out-of-sample forecast errors than on maximising in-sample 'goodness of fit'. Thus, the approach followed is unashamedly one of 'model mining' with the aim of optimising forecast performance.[4]

The structure of the paper is as follows: Section 2 presents a brief introduction to ARIMA modelling, outlining the main advantages and disadvantages. Section 3 - the main section of the paper - outlines a general framework for ARIMA forecasting, including a comparison of the traditional Box-Jenkins methodology with objective penalty function methods. A practical application of the framework is made with reference to the HICP series and its major sub-components. The discussion focuses on two series - the overall HICP and the non-energy industrial goods (HICPC) component - as these serve to highlight many of the issues encountered when using ARIMA models to forecast inflation. Section 4 briefly summarises a semi-automatic algorithm developed in the preparation of this paper. Section 5 concludes and offers some observations on the limitations of ARIMA models. Appendix A provides a

---

[3] For information on the construction of an historical series for the HICP in Ireland see Meyler *et al* (1998).

4 Cecchetti (1995, pg. 199) finds that, in his study, "whether a model fits well in-sample tells us virtually nothing about its out-of-sample forecasting ability". However, in this paper, a positive correlation is generally found between a model's in-sample explanatory power as ranked according to the penalty function criterion and its out-of-sample forecast rank according to the sum of the average mean absolute error for each of the first four steps ahead.

description of ARIMA models and some of their theoretical properties. Appendices B and C present results for the other sub-components of the HICP.

2. An Introduction to ARIMA Modelling

ARIMA methods for forecasting time series are essentially agnostic. Unlike other methods they do not assume knowledge of any underlying economic model or structural relationships. It is assumed that past values of the series plus previous error terms contain information for the purposes of forecasting.

The main advantage of ARIMA forecasting is that it requires data on the time series in question only. First, this feature is advantageous if one is forecasting a large number of time series. Second, this avoids a problem that occurs sometimes with multivariate models. For example, consider a model including wages, prices and money. It is possible that a consistent money series is only available for a shorter period of time than the other two series, restricting the time period over which the model can be estimated. Third, with multivariate models, timeliness of data can be a problem. If one constructs a large structural model containing variables which are only published with a long lag, such as wage data, then forecasts using this model are conditional forecasts based on forecasts of the unavailable observations, adding an additional source of forecast uncertainty.

Some disadvantages of ARIMA forecasting are that:

- Some of the traditional model identification techniques are subjective and the reliability of the chosen model can depend on the skill and experience of the forecaster (although this criticism often applies to other modelling approaches as well).
- It is not embedded within any underlying theoretical model or structural relationships. The economic significance of the chosen model is therefore not clear. Furthermore, it is not possible to run policy simulations with ARIMA models, unlike with structural models.[5]

[5] For a discussion of this issue see Frain (pg. 12, 1995).

- ARIMA models are essentially 'backward looking'. As such, they are generally poor at predicting turning points, unless the turning point represents a return to a long-run equilibrium.

However, ARIMA models have proven themselves to be relatively robust especially when generating short-run inflation forecasts. ARIMA models frequently outperform more sophisticated structural models in terms of short-run forecasting ability (see, for example, Stockton and Glassman (1987) and Litterman (1986)). Therefore, the ARIMA forecasting technique outlined in this paper will not only provide a benchmark by which other forecasting techniques may be appraised, but will also provide an input into forecasting in its own right.

Appendix A presents a description of ARIMA models and some of their theoretical properties. A general notation for a multiplicative seasonal ARIMA models is ARIMA (p,d,q)(P,D,Q), where p denotes the number of autoregressive terms, q denotes the number of moving average terms and d denotes the number of times a series must be differenced to induce stationarity. P denotes the number of seasonal autoregressive components, Q denotes the number of seasonal moving average terms and D denotes the number of seasonal differences required to induce stationarity.

This may be written as

(1) $$\phi(B)\Phi(B)\nabla^{d}\nabla_{s}^{D}Y_{t} = \theta(B)\Theta(B)a_{t}$$

where,

$X_t = \nabla^{d}\nabla_{s}^{D}Y_{t}$ is a stationary series,

$\nabla^{d} = (1-B)^{d}$ represents the number of regular differences and $\nabla_{s}^{D} = (1-B^{s})^{D}$ represents the number of seasonal differences required to induce stationarity in Y_t,

s is the seasonal span (hence for quarterly data s = 4 and for monthly data s = 12),

B is the backshift operator (such that $B^{0}X_{t} = X_{t}$, $B^{1}X_{t} = X_{t-1}$, $B^{2}X_{t} = X_{t-2}$, ..),

$\theta(B) = 1+\theta_{1}B+\theta_{2}B^{2}+...+\theta_{q}B^{q}$ is a q-order polynomial in the backshift operator,

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$,

$\Phi(B) = 1 - \Phi_{1s} B^{1s} - \Phi_{2s} B^{2s} - .. - \Phi_{Ps} B^{Ps}$, and

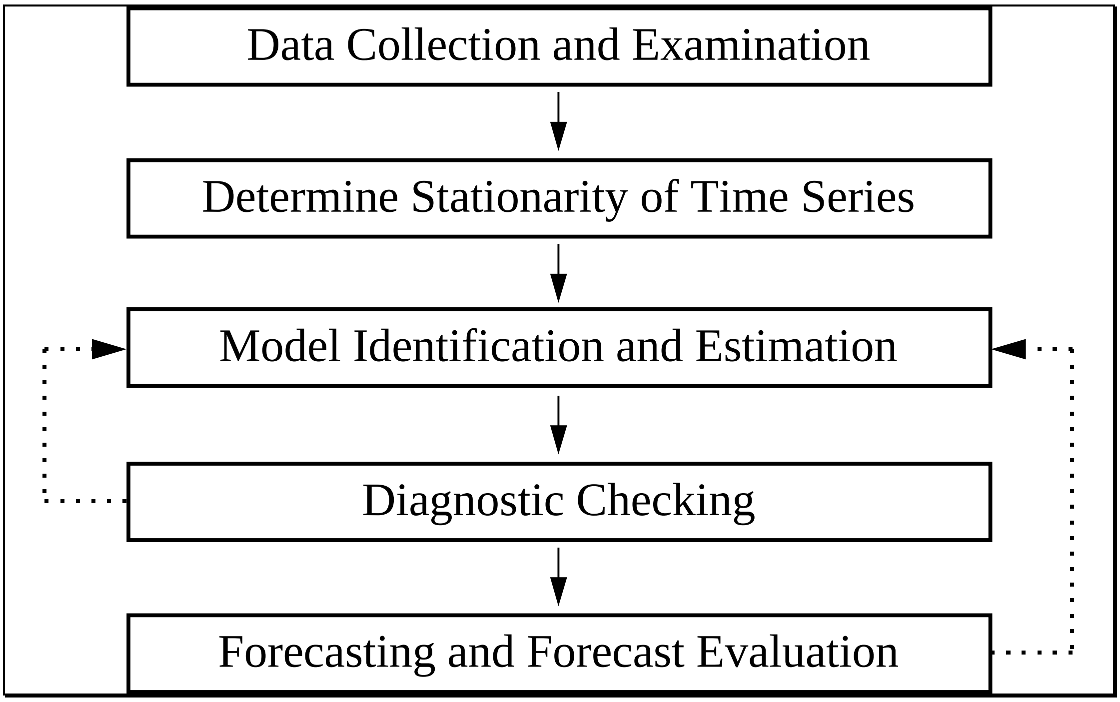
$\Theta(B) = 1 + \Theta_{1s} B^{1s} + \Theta_{2s} B^{2s} + .. + \Theta_{Qs} B^{Qs}$,

As shown in Appendix A any non-deterministic stationary process can be approximated by an ARMA process. **The problem lies in ensuring the series is stationary and in determining the order of p and q that adequately describes the time series being examined.** It is these issues which are examined in the next section.

3. ARIMA FORECASTING IN PRACTICE

This section outlines a general ARIMA modelling and forecasting strategy. Figure 1 illustrates this process graphically. It is important to note, however, that this process is not a simple sequential one, but can involve iterative loops depending on results obtained at the diagnostic and forecasting stages. The first step is to collect and examine graphically and statistically the data to be forecast. The second step is to test whether the data are stationary or if differencing is required. Once the data are rendered stationary one should seek to identify and estimate the correct ARMA model. Two alternative approaches to model identification are considered - the Box-Jenkins methodology and penalty function criteria. It is important that any identified model be subject to a battery of diagnostic checks (usually based on checking the residuals) and sensitivity analysis. For example, the estimated parameters should be relatively robust with respect to the time frame chosen. Should the diagnostic checks indicate problems with the identified model one should return to the model identification stage. Once a model or selection of models has been chosen, the models should then be used to forecast the time series, preferably using out-of-sample data to evaluate the forecasting performance of the model. One common pitfall of ARIMA modelling is to overfit the model at the identification stage, which maximises the in-sample explanatory performance of the model but may lead to poor out-of-sample predictive power relative to a more parsimonious model. Thus, if a model with a large number of AR and MA lags yields poor forecasting performance, it may be optimal to return to the model identification stage and consider a more parsimonious model.

FIGURE 1 - ARIMA FORECASTING PROCEDURE



3.1. STEP ONE - DATA COLLECTION AND EXAMINATION

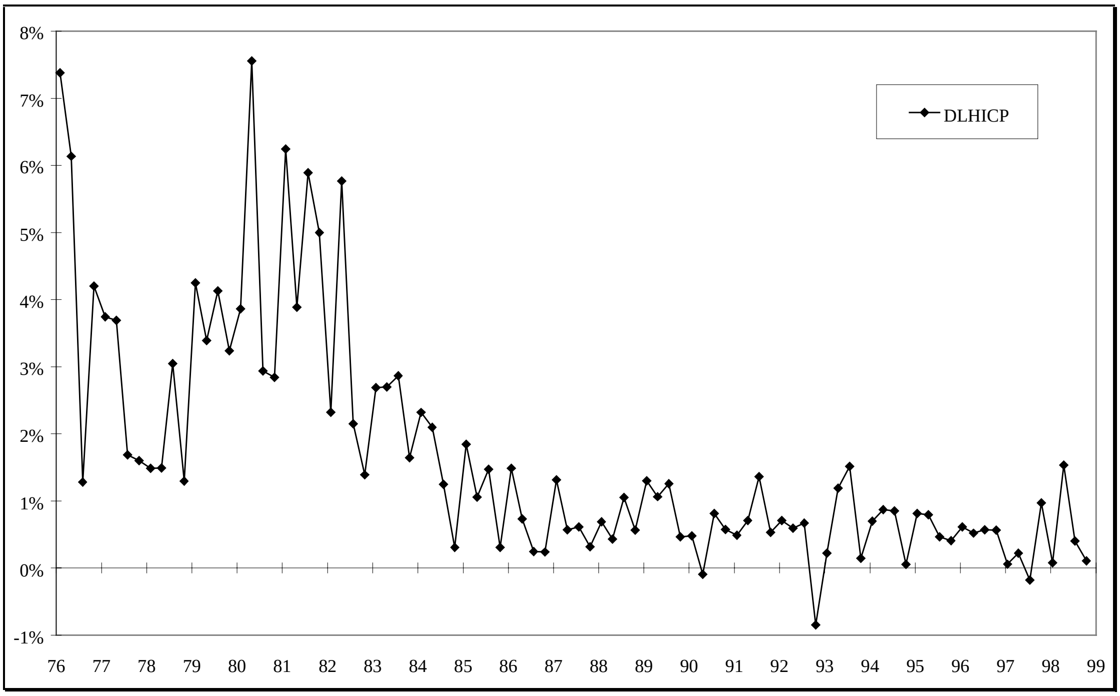
*"An econometrician should **always** fall in love with his/her **data**"*

A lengthy time series of data is required for univariate time series forecasting. It is usually recommended that at least 50 observations be available. Using either Box-Jenkins or objective penalty function methods can be problematic if too few observations are available. Unfortunately, even if a long time series is available, it is possible that the series contains a structural break which may necessitate only examining a sub-section of the entire data series, or alternatively using intervention analysis or dummy variables. Thus, there may be some conflict between the need for sufficient degrees of freedom for statistical robustness and having a shorter data sample to avoid structural breaks.

Graphically examining the data is important. They should be examined in levels, logs, differences and seasonal differences. The series should be plotted against time to assess whether any structural breaks, outliers or data errors occur. If so one may need to consider use of intervention or dummy variables. This step may also reveal whether there is a significant seasonal pattern in the time series.

Consider, for example, a plot of the first difference of the log of the HICP series for the period Q1 1976 to Q4 1998 as shown in Figure 2. From this figure and Table 1 it is evident that for the period 1976 to 1983, the mean rate of, and standard deviation of, inflation was higher than for the period post-1983. Thus, it may be necessary to consider inclusion of an intervention variable for the earlier period, or perhaps, to identify and estimate the model for the later period only.[6]

FIGURE 2 - PLOT OF DLHICP[7] SERIES, 1976Q1-1998Q4



---

[6] A more formal test (Perron, 1989) for nonstationarity, in the presence of structural breaks, is considered below.

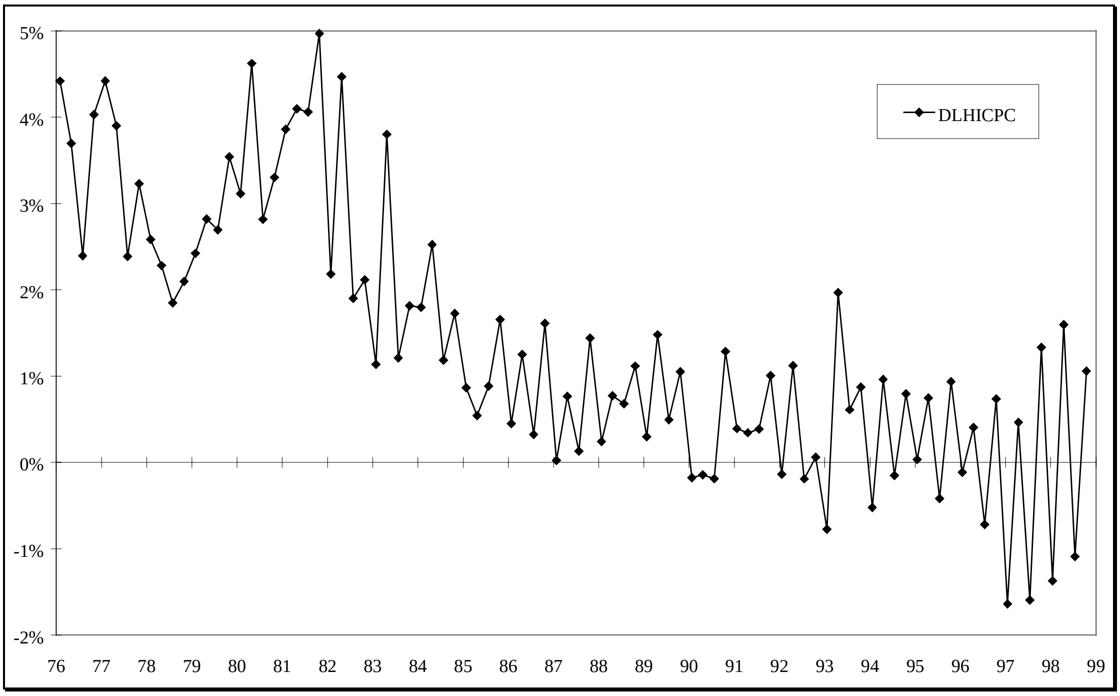
[7] The following notation is used in this paper. LHICP denotes the natural log of the HICP series. DLHICP denotes the LHICP series differenced once. DDLHICP denotes the LHICP series differenced twice. DsDLHICP denotes the seasonal difference of the DLHICP series.

TABLE 1 - SUMMARY STATISTICS FOR DLHICP SERIES, 1976Q1-1998Q4

Period	Mean	Standard Deviation
1976Q1-1983Q4	3.49	1.79
1984Q1-1998Q4	0.71	0.56
Overall Period	1.68	1.75

A plot of the first difference of the log of the non-energy industrial goods component of the HICP (HICPC), given in Figure 3, also yields useful information. This series shows a decline in the mean rate of change similar to that for the DLHICP series. Another noticeable feature is the more violent oscillation in recent periods. This appears to reflect a more pronounced seasonal pattern with deeper sales discounts and subsequent rebound in prices. The more pronounced seasonal sales pattern in the HICPC index is considered further below.

FIGURE 3 - PLOT OF DLHICPC SERIES, 1976Q1-1998Q4



Another way to examine the properties of a time series is to plot its autocorrelogram. The autocorrelogram plots the autocorrelation between differing lag lengths of the time series. Plotting the autocorrelogram is a useful aid for determining the stationarity of a time series, and is also an important input into Box-Jenkins model identification. The theoretical autocorrelogram for different orders of AR, MA and ARMA models are outlined in section dealing with model identification (Step 3). The SACF may be constructed using equation (A4). The maximum lag length considered is usually no more than n/4. Although sample autocorrelations for lags in excess of twice the seasonal span (i.e., in excess of 8 for quarterly data) should be treated with caution.

If a time series is stationary then its autocorrelogram should decay quite rapidly from its initial value of unity at zero lag. If the time series is nonstationary then the autocorrelogram will only die out gradually over time. Figure 4 plots the autocorrelogram for the log of the HICP, the first differences and second differences of the log and the seasonal difference of the first difference (for the period 1984Q1 - 1998Q4). It would appear from Figure 4 that the log of levels series is nonstationary as the autocorrelations decay slowly towards zero. At first glance the first difference series appears stationary although there seems to be some evidence of seasonal behaviour (the autocorrelations at lags 4, 8 and 12 exhibit distinctive behaviour and die out quite slowly). The autocorrelogram of the second difference series is more volatile and may indicate over-differencing. The autocorrelations of the seasonal differences of the first difference series exhibit a quasi-sinusoidal decay pattern, which may indicate the presence of complex roots. Based on a graphical examination of Figure 4, the first difference of logs and the seasonal differences of the first differences require more formal unit root testing to determine stationarity.

FIGURE 4 -

AUTOCORRELOGRAM FOR LHICP, DLHICP, DDLHICP AND DSDLHICP

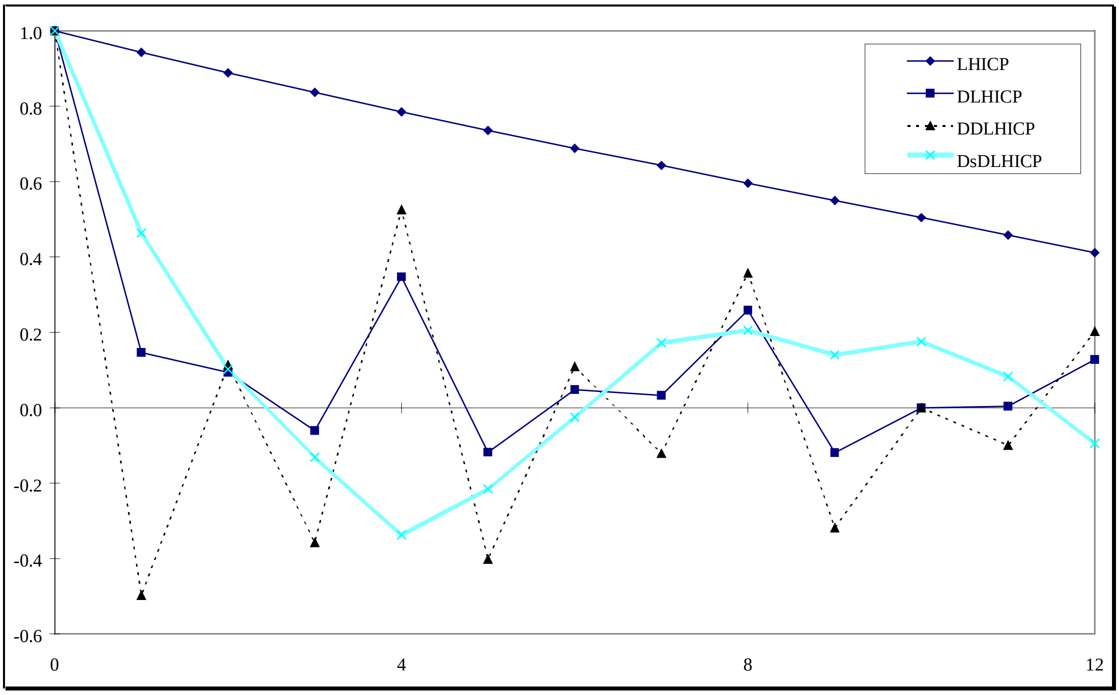
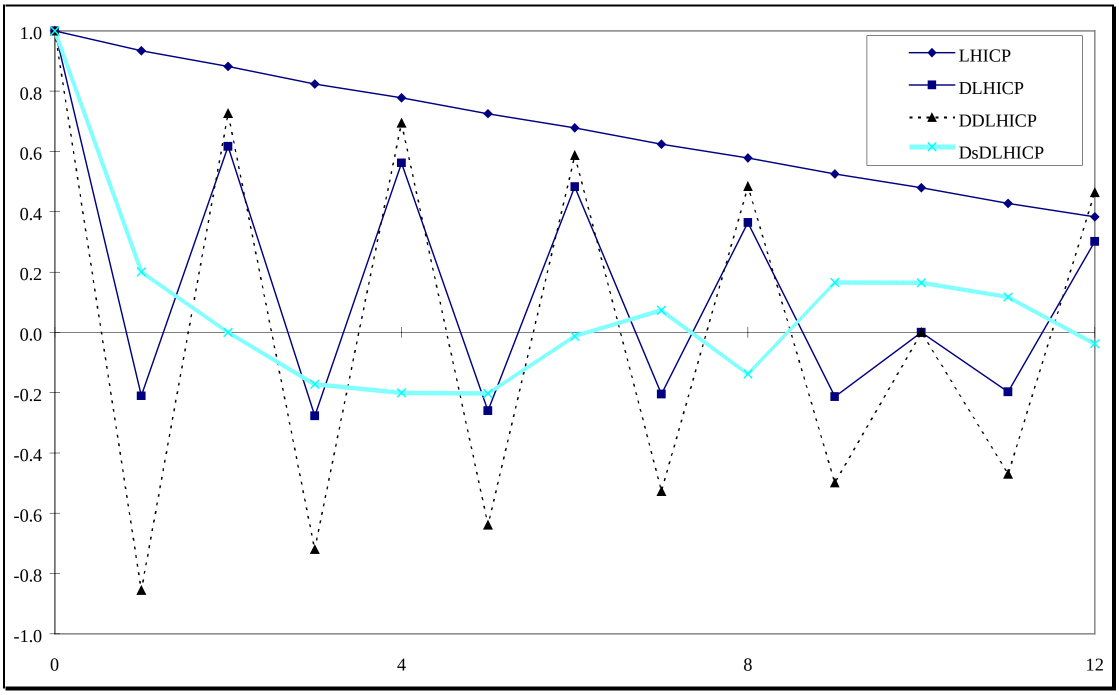


Figure 5 displays the autocorrelograms for various transformations of the non-energy industrial goods (HICPC) component of the HICP. It is evident that the goods component of the HICP exhibits a much stronger seasonal pattern than the overall HICP. This pattern is driven primarily by the Winter and Summer sales and the subsequent rebound. Based on Figure 5 there is a strong case for seasonally differencing the rate of inflation in non-energy industrial goods prices (i.e., seasonally differencing the first difference of the log of the HICPC series).

FIGURE 5 -

AUTOCORRELOGRAM FOR LHICPC, DLHICPC, DDLHICPC AND DSDLHICPC



Although the autocorrelogram gives some indication as to whether a series is stationary or nonstationary, in more recent years a vast array of formal tests for stationarity with known statistical properties have been developed.

3.2. STEP TWO - TESTING FOR STATIONARITY

The time series under consideration must be stationary before one can attempt to identify a suitable ARMA model. A large literature has developed in recent years on the issue of testing time series for stationarity and nonstationarity (See, for example, Harris (1995) and Banerjee *et al* (1993)).

For AR or ARMA models to be stationary it is necessary that the modulus of the roots of the AR polynomial be greater than unity, and for the MA part to be invertible it is also necessary that the roots of the MA polynomial lie outside the unit circle.

The original Dickey-Fuller test considered the model $X_t = \rho X_{t-1} + \varepsilon_t$ or alternatively $\Delta X_t = (\rho-1)X_t + \varepsilon_t$. If the series contains a unit root, then $\rho = 1$ (and $\rho-1 = 0$). The standard t-distribution cannot be used to test if $\rho = 1$, the Dickey-Fuller distribution should be used instead. However, should ε_t be autocorrelated, the Dickey-Fuller distribution is no longer valid either. In this case, an alternative model should be estimated, where *l* lags of the first difference of the series are added until the series e_t displays no evidence of autocorrelation.[8]

(2) $$\Delta X_t = (\rho-1)X_t + \sum_{i=1}^{l} \delta_i \Delta X_{t-i} + e_t$$

In this instance the Augmented Dickey Fuller test statistic should be used. Table 2 presents some summary results testing the HICP series and its major sub-components for unit roots. The t-adf statistic is the Augmented Dicky Fuller test statistic, under the null hypothesis that $\rho-1 = 0$ (or equivalently, $\rho = 1$). The columns denoted lags indicate the number of lagged first differences of the series that were added to ensure 'white noise' error terms. The final column in each part of the table contains the 5 per cent and 1 per cent critical values for the t-adf statistic. For the log of the levels series, the series is tested for nonstationarity around a constant and a trend. However, the differenced series are tested for nonstationarity around a constant solely. The results from Table 2 indicate that: none of the levels series, except HICPA and HICPB, is stationary around a constant and trend; all of the series except the HICPB and HICPC series are stationary if differenced once or more.[9] However, the t-adf statistics on the HICP and HICPE series differenced once are relatively low, and may indicate the need to consider seasonal differencing.[10] This issue is considered further below.

---

[8] In practice, the number of lag differenced terms to be added is determined using model selection criteria such as the AIC and BIC outlined below. This is usually sufficient to ensure a well-behaved error term.

[9] In general it is sufficient to test up to d = 2 and D = 1.

[10] The conflicting results for the HICPB series and the result for the HICPA series in levels indicate the low power of many unit root tests. This is why it is necessary to use a number of alternative tests to ensure consistency.

TABLE 2 - AUGMENTED DICKY-FULLER TESTS, 1984Q1 - 1998Q4

	HICP		HICPA		HICPB		5% (1%)
	t-adf	lags	t-adf	lags	t-adf	lags	t-adf*
(1) Log of Level	-3.1	4	-4.1*	3	-4.3**	4	-3.5 (-4.1)
(2) 1st Diff. of (1)	-3.8**	4	-7.5**	1	-2.3	3	-2.9 (-3.5)
(3) Seas. Diff. of (2)	-5.1**	3	-10.1**	3	-6.6**	3	-2.9 (-3.5)
(4) 2nd Diff. of (2)	-5.0**	4	-11.6**	1	-9.2**	1	-2.9 (-3.5)

	HICPC		HICPD		HICPE		5% (1%)
	t-adf	lags	t-adf	lags	t-adf	lags	t-adf*
(1) Log of Level	-2.0	4	-3.5	4	-1.6	4	-3.5 (-4.1)
(2) 1st Diff. of (1)	-2.1	3	-7.8**	0	-4.3**	3	-2.9 (-3.5)
(3) Seas. Diff. of (2)	-6.2**	0	-5.7**	3	-3.5**	3	-2.9 (-3.5)
(4) 2nd Diff. of (2)	-6.0**	4	-8.3**	2	-11.5**	2	-2.9 (-3.5)

A number of less formal techniques exist for determining stationarity of a time series. As stated above examination of the autocorrelogram can be one useful indicator. For example, consider a pure AR(1) process. The autocorrelation at lag k of an AR(1) process is given by ϕ_1^k. Thus if $\phi_1 = 1$ the autocorrelogram does not decay over time. In general, the sample autocorrelogram of a nonstationary series will only decay very slowly towards zero. However, relying solely on the autocorrelogram to determine the stationarity of a time series tends to lead to over-differencing (Mills (1990, pg. 121)). One way to check for over-differencing is to examine the variance of the process (Anderson, 1976). In general the sample variance of a process will decrease until the correct order of differencing is found, but will increase thereafter if the process is over-differenced. Table 3 indicates that using log of levels is inappropriate for all the HICP series. The sample variances for the HICPC and HICPE series suggest seasonally differencing the inflation rate could be necessary as the inflation rate itself may be nonstationary. For the other series (HICP, HICPA, HICPB and HICPD), the sample variances indicate that using the inflation rate is sufficient to ensure stationarity, although the results are not always clear-cut.

TABLE 3 - SAMPLE VARIANCE, 1984Q1 - 1998Q4

		HICP	HICPA	HICPB	HICPC	HICPD	HICPE
(1)	Log of Level	0.01280	0.00591	0.01745	0.00715	0.00222	0.02886
(2)	1st Diff. of (1)	**0.00003**	**0.00047**	**0.00009**	0.00007	**0.00038**	0.00012
(3)	Seas. Diff. of (2)	0.00004	0.00106	0.00010	**0.00004**	0.00068	**0.00010**
(4)	2nd Diff. of (2)	0.00005	0.00093	0.00017	0.00018	0.00077	0.00038

In summary, whilst the adf test and analysis of the sample variance favour a single differencing of the overall HICP to induce stationarity, examination of the autocorrelogram suggests seasonal differencing. Furthermore, the variance of the seasonal difference of the first difference of the LHICP series is only marginally higher than the LHICP series just differenced once. In contrast, the first difference of the HICPC series fails the adf test at the 5 per cent confidence level. The seasonal difference of the first difference of the HICPC series passes all the tests for stationarity. In general, when different test statistics offer conflicting evidence, it is best to bring forward both alternatives to the forecasting stage, as the power of many tests for nonstationarity can be quite low especially with small sample sizes.

3.3. STEP THREE - MODEL IDENTIFICATION AND ESTIMATION

Having determined the correct order of differencing required to render the series stationary, the next step is to find an appropriate ARMA form to model the stationary series. There are two main approaches to identification of ARMA models in the literature. The traditional method utilises the Box-Jenkins procedure, in which an iterative process of model identification, model estimation and model evaluation is followed. The Box-Jenkins procedure is a quasi-formal approach with model identification relying on subjective assessment of plots of autocorrelograms and partial autocorrelograms of the series. Objective measures of model suitability, in particular the penalty function criteria, have been used by some authors instead of the traditional Box-Jenkins procedure. For a recent example of the use of objective penalty function criteria see Gómez and Maravall (1998). However, these ‘objective’ measures are not without problems either.

Outside of the Box-Jenkins and penalty function criterion methods there are a number of alternative identification methods proposed in the literature. These include, *inter alia*, the Corner method (Beguin *et al*, 1980), the R and S Array method (Gray *et al*, 1978), and canonical correlation methods (Tsay and Tiao, 1985). These methods are usually based on the properties of the autocorrelation function and do not require estimation of a range of models, which can be computationally expensive. This lack of computation is a significant advantage over the penalty function criterion outlined below. However, the problem with most autocorrelation-based methods is that they are not very useful for dealing with seasonal data. The seasonal nature of price data makes these alternative methods less attractive for the purposes of forecasting inflation.

3.3.1. Box-Jenkins Methodology

The Box-Jenkins methodology essentially involves examining plots of the sample autocorrelogram, partial autocorrelogram and inverse autocorrelogram and inferring from patterns observed in these functions the correct form of ARMA model to select.[11] The Box-Jenkins methodology is not only about model identification but is, in fact, an iterative approach incorporating model estimation and diagnostic checking in addition to model identification.

Theoretically Box-Jenkins model identification is relatively easy if one has a pure AR or a pure MA process. However, in the case of mixed ARMA models (especially of high order) it can be difficult to interpret sample ACFs and PACFs, and Box-Jenkins identification becomes a highly subjective exercise depending on the skill and experience of the forecaster. Random noise in time series, especially price data, makes Box-Jenkins model identification even more problematic.

Pure AR Process

The autocorrelations of a pure AR(p) process should decay gradually at increasing lag length. Hence, using an autocorrelogram it is not possible to differentiate between a pure AR(3) model or a pure AR(4) model. However, the partial autocorrelations of a pure AR(p) process do display distinctive features. The partial autocorrelogram should 'die out' after p lags. Thus, the partial autocorrelogram of a pure AR(3) process should die out after 3 lags, whereas that of a pure AR(4) process would die out after 4 lags.

Hence, for a pure AR(p) process the theoretical ACF and PACF are as follows:

$ACF(i) \neq 0 \ \forall \ i$
$PACF(i) \neq 0 \ \forall \ i = 1,...,p$
$PACF(i) = 0 \ \forall \ i > p$

where i denotes the number of lags.

Pure MA Process

The behaviour of correlograms and partial autocorrelograms for pure MA(q) processes is the reverse of that for pure AR processes. The autocorrelogram of a pure MA(q) process should 'die out' after q lags. The partial autocorrelogram of a pure MA process, on the other hand, only decays slowly over time (similar to the behaviour of the autocorrelogram of a pure AR process). Thus, it should be impossible to distinguish between the PACF of an MA(3) and MA(4) process, whereas the ACF of the MA(3) process should decay to zero after 3 lags and the MA(4) process after 4 lags.

---

[11] If the ARMA model $\phi(B)X_t = \theta(B)a_t$ is invertible, the inverse autocorrelogram of the series, X_t, is simply the autocorrelogram of the inverted model (i.e., the 'dual' of the original model) given by $\theta(B)X_t = \phi(B)a_t$. See, Chatfield (1979) for a discussion of inverse autocorrelation functions.

Hence, for a pure MA(q) process the theoretical ACF and PACF are as follows:

ACF(i) $\neq 0 \; \forall \; i = 1,...,q$
ACF(i) $= 0 \; \forall \; i > q$
PACF(i) $\neq 0 \; \forall \; i$

Thus if one has either a pure AR or MA process model identification should be relatively straightforward in theory. Furthermore the behaviour of the autocorrelogram and partial autocorrelogram can provide information on the AR and MA components, in terms of sign or the existence of complex roots. For example, the autocorrelations of a pure AR(1) process with a negative root should oscillate around zero and decay with increases in lags, whereas the autocorrelations of a pure AR(1) process with a positive root should decay gradually and monotonically towards zero (assuming $|\phi_1| \leq 1$). The autocorrelogram of an AR(p) process with complex roots should exhibit a sinusoidal (or wave) pattern.

Mixed ARMA Processes

Unfortunately, model identification is greatly complicated for mixed (i.e., ARMA) processes. The patterns of sample autocorrelations and partial autocorrelations of high order ARMA models are notoriously difficult to interpret. Thus, model identification using the Box-Jenkins procedures will be an iterative process, with Step Four - diagnostic checking - determining whether alternative models should be examined. See Box and Jenkins (1976) for a detailed discussion of identifying mixed ARMA process.

3.3.2. Objective Model Identification

Because of the highly subjective nature of the Box-Jenkins methodology, time series analysts have sought alternative objective methods for identifying ARMA models. Penalty function statistics, such as Akaike Information Criterion [AIC] or Final Prediction Error [FPE] Criterion (Akaike, 1974), Schwarz Criterion [SC] or Bayesian Information Criterion [BIC] (Schwarz, 1978) and Hannan Quinn Criterion [HQC]

(Hannan, 1980), have been used to assist time series analysts in reconciling the need to minimise errors with the conflicting desire for model parsimony. These statistics all take the form minimising the sum of the residual sum of squares plus a 'penalty' term which incorporates the number of estimated parameter coefficients to factor in model parsimony.

These statistics take the form

(3) $$\text{BIC} = \log\left(\frac{rss}{n}\right) + \left(\log(n) * \frac{k}{n}\right),$$

(4) $$\text{HQC} = \log\left(\frac{rss}{n}\right) + \left(2 * \log(\log(n)) * \frac{k}{n}\right), \text{ and}$$

(5) $$\text{AIC} = \log\left(\frac{rss}{n}\right) + \left(2 * \frac{k}{n}\right)$$

where,
k = number of coefficients estimated (1 + p + q + P + Q)
rss = residual sum of squares
n = number of observations.

Assuming there is a true ARMA model for the time series, the BIC and HQC have the best theoretical properties. The BIC is strongly consistent whereas AIC will usually result in an overparameterised model; that is a model with too many AR or MA terms (Mills 1993, p.29). Indeed, it is easy to verify that for n greater than seven the BIC imposes a greater penalty for additional parameters than does the AIC. Gómez and Maravall (1998, p.19) also favour the BIC over the AIC.

Thus, in practice, using the objective model selection criteria involves estimating a range of models and the one with the lowest information criterion is selected. This can create a number of difficulties. First, it can be computationally expensive using the penalty function criterion. Estimating all possible models encompassed by a

(3,0,3)(2,0,2) model involves estimating 144 different models. Therefore the choice of maximum order is very important to avoid expensive computational requirements. Unfortunately, there is no *a priori* information to assist in selecting the maximum order of the ARIMA model to estimated. Moving from a maximum order of (3,0,3)(2,0,2) to (2,0,2)(1,0,1) reduces the number of models to be estimated from 144 to 36. One useful rule of thumb for determining the maximum order is to select a maximum for the regular terms of the seasonal span less one (i.e., three for quarterly data or eleven for monthly data) and one for the seasonal term. Thus for quarterly data this would suggest estimating an ARMA of maximum order (3,0,3)(1,0,1), which implies estimating 64 different models and calculating 64 information criterion.[12]

Second, the different objective model selection criteria can suggest different models. That is the ranking order based on the BIC will usually not be the same as under the AIC. Table 4 compares the top five ranking models under the BIC, HQC and AIC for the DLHICP series estimated over the period 1984Q1 - 1998Q4. The top ranking model under the AIC only ranks seventh using the BIC. Furthermore, the AIC generally favours a less parsimonious model than either the BIC or the HQC.

TABLE 4 - COMPARISON OF RANKING BY CRITERION - DLHICP (1984Q1 - 1998Q4)

	BIC		HQC		AIC	
rank 1	(0,0,0) x (1,0,1)	-10.646	(0,0,0) x (1,0,1)	-10.709	(3,0,0) x (1,0,0)	-10.773
rank 2	(0,0,0) x (1,0,0)	-10.637	(3,0,0) x (1,0,0)	-10.705	(3,0,0) x (1,0,1)	-10.772
rank 3	(1,0,0) x (1,0,0)	-10.629	(1,0,0) x (1,0,0)	-10.693	(0,0,0) x (1,0,1)	-10.750
rank 4	(0,0,1) x (1,0,0)	-10.618	(1,0,0) x (1,0,1)	-10.692	(1,0,0) x (1,0,1)	-10.747
rank 5	(1,0,0) x (1,0,1)	-10.607	(3,0,0) x (1,0,1)	-10.690	(0,0,1) x (1,0,1)	-10.738

Third, even if one utilises only one measure (e.g., BIC), the difference between the BIC statistic for different models is sometimes only marginal. Poskitt and Tremayne (1987) suggest the idea of a model portfolio. This involves comparing alternative

[12] In addition to the MA and AR dimensions, it is also necessary to determine the correct level of differencing. For the analysis in this paper, all the series were also seasonally differenced in addition to a single regular differencing and fitted with an ARMA model.

models to the best model suggested by the information criterion. Denoting the best model as (p*,0,q*)(P*,0,Q*), the following statistic is computed for each alternative model

$$(6) \qquad \Re = \exp\left[-\frac{1}{2}T\left\{BIC\big((p^*,0,q^*)(P^*,0,Q^*)\big) - BIC\big((p,0,q)(P,0,Q)\big)\right\}\right].$$

Although this statistic has no formally defined critical values, it may be used to quantify the decisiveness with which a particular model can be rejected compared to the 'best' model. Poskitt and Tremayne suggest that a value of less than $\sqrt{10}$ as a suitable cut-off point. This implies that models with an information criterion within $2\log\left(\sqrt{10}\right)/T$ of the best model enter the model portfolio. Then, using the model portfolio approach, not only the 'best' model is used in the diagnostic and forecasting stages but all models in the portfolio. Using the DLHICP series over the period 1984Q1-1998Q4 (T = 60), the portfolio approach would consider three extra models in addition to the model which minimised the BIC statistic. The largest portfolio using any of the six series considered in this paper would contain five models.[13]

Gómez and Maravall (1998, p.21) suggest using balanced models where possible if two models perform relatively similarly. In other words, select a model where p and q are relatively similar rather than choosing a model with just AR terms or just MA terms. For example, if the information criteria suggest that a (2,0,0) and a (1,0,1) model perform similarly then Gómez and Maravall would suggest using the more balanced model (1,0,1). One benefit of choosing balanced models might be that it would be easier to identify common factors in the AR and MA polynomials. If an ARIMA model has common factors it should be possible to represent the model in a more parsimonious manner by eliminating the common factor from both the AR and MA components.

---

[13] This was the DLHICPA series. Using the critical value above, only one model would enter the portfolio for the DLHICPB, DLHICPC and DLHICPE series.

A more general alternative approach to the model portfolio approach outlined above is simply to select the top five or ten ranking models and carry these forward to the diagnostic checking and forecasting stage. This is the approach adopted in the semi-automatic ARIMA model selection algorithm outlined below.

Table 5 below presents the top five ranking models as classified by the BIC for the DLHICP, DLHICPC and DsDLHICPC series. The models indicated in bold type are those models which would enter the model portfolio using the criterion of $\sqrt{10}$ suggested by Poskitt and Tremayne (1987).[14] The results are shown for the regularly differenced (DLHICPC) and the regularly differenced plus seasonally differenced non-energy industrial goods series (DsDLHICPC) as the tests for nonstationarity were indeterminate between a single regular differencing and a seasonal differencing in addition to the regular differencing. The BIC indicates that the regularly differenced series performs slightly better than the seasonally differenced series, reflecting the ambiguity over the correct order of differencing.

TABLE 5 - TOP FIVE MODELS BASED ON BIC -
DLHICP, DLHICPC AND DSDLHICPC (1984Q1 - 1998Q4)

Rank	**DLHICP**		**DLHICPC**		**DLHICPC**	
	Model	**BIC**	**Model**	**BIC**	**Model**	**BIC**
1	**(0,0,0) x (1,0,1)**	**-10.646**	**(0,0,0) x (1,0,0)**	**-10.162**	**(0,0,0) x (1,1,1)**	**-10.062**
2	**(0,0,0) x (1,0,0)**	**-10.637**	(0,0,0) x (1,0,1)	-10.104	**(0,0,0) x (0,1,1)**	**-10.045**
3	**(1,0,0) x (1,0,0)**	**-10.629**	(1,0,0) x (1,0,0)	-10.094	**(0,0,0) x (0,1,0)**	**-10.038**
4	**(0,0,1) x (1,0,0)**	**-10.618**	(0,0,1) x (1,0,0)	-10.094	(0,0,0) x (1,1,0)	-10.016
5	(1,0,0) x (1,0,1)	-10.607	(3,0,0) x (1,0,0)	-10.074	(1,0,0) x (1,1,1)	-10.010

In summary, the main advantages and disadvantages of objective penalty function criteria are as follows:

Advantages of Objective Penalty Function Criteria

- Objective measure with no subjective interpretation.
- Results are readily reproducible and verifiable.
- BIC and HQC are asymptotically consistent.

Disadvantages of Objective Penalty Function Criteria

- Need to calculate a wide range of models. This can be computationally expensive.
- There are no theoretical guidelines for choosing the maximum order of ARIMA model to consider.
- Sometimes there is little to chose between competing models.

3.4. STEP FOUR - MODEL DIAGNOSTICS

"An econometrician should ***never*** *fall in love with his/her* ***model****"*

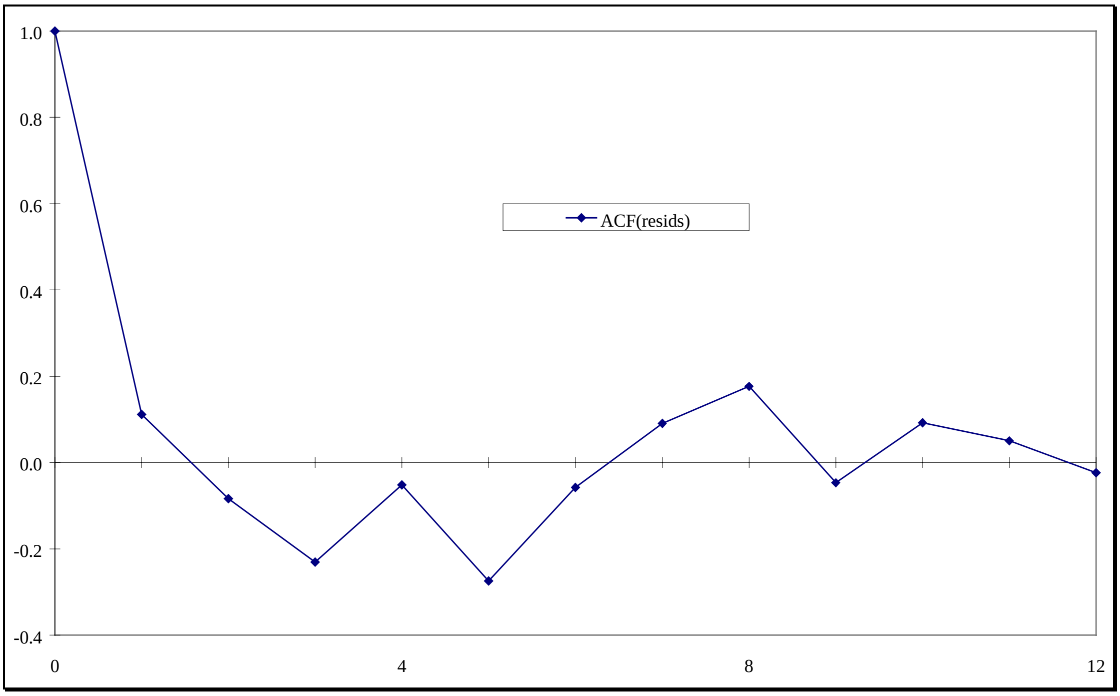
The fourth step will be the formal assessment of each of the time series models. This will involve a rigorous assessment of the diagnostic tests for each of the competing models. As different models may perform reasonably similarly, a number of alternative formulations may have to be retained at this stage to be further assessed at the forecasting stage.

There are a number of diagnostic tools available for ensuring a satisfactory model is arrived at. Plotting the residuals of the estimated model is a useful diagnostic check. This should indicate any outliers that may affect parameter estimates and also point towards any possible autocorrelation or heteroscedacity problems. A second check of model suitability is to plot the autocorrelogram of the residuals. If the model is correctly specified the residuals should be 'white noise'. Therefore, a plot of the autocorrelogram should immediately die out from one lag on. Any significant autocorrelations may indicate that the model is misspecified and may point to the solution. For example, if a (0,0,1)(0,0,0) model of a quarterly time series is estimated,

[14] Note, however, that an indication of the relative stability of the DLHICP series over the period (1984Q1 - 1998Q4) is that fitting a straight line (i.e., a (0,0,0)(0,0,0) model) to the series ranks 21st out of 64 models under the BIC (-10.327).

but the autocorrelogram of the residuals indicates a significant autocorrelation at the fourth lag, this would indicate that a (0,0,1)(0,0,1) model should be estimated, as this might remove the autocorrelation at the fourth (seasonal) lag. Figure 6 plots the ACF of the residuals from the (0,0,0)(1,0,1) model of DLHICP fitted over the period 1984Q1 - 1998Q4. In general, the autocorrelations are not significantly different from zero, however, the autocorrelations at lags three, five and eight are marginally significant.

FIGURE 6 - PLOT OF ACF OF RESIDUALS FROM (0,0,0)(1,0,1) MODEL OF DLHICP SERIES, 1984Q1-1998Q4



More formal test statistics exist which involve testing the residuals of the estimated model. The Ljung-Box (1978) Q statistic is the most commonly used test statistic. The Q-statistic tests for autocorrelation in the residuals where Q is defined as

$$(7) \qquad Q(k) = T(T+2)\sum_{i=1}^{k}(T-i)^{-1} r_i^2 \ \sim^{a} \ \chi_k^2 .$$

Another essential check is to test the robustness of a selected model by estimating it over a number of different time periods. If the parameter estimates are not stable over time this indicates that further consideration will have to be given to the model. It may be that the time series contains a structural break and that, for the purposes of forecasting, only the period since the structural break should be used when estimating the model, as there can be a fundamental conflict between estimating a model to maximise in-sample goodness-of-fit and optimising out-of-sample forecast performance when the series contains a structural break. A formal test for nonstationarity in the presence of a structural break may be carried out using Perron's (1989) augmented unit root test. Examining the DLHICP data using Perron's augmented unit root test, allowing for a structural break, indicates that the data are stationary post-1983, with the statistic on the structural break being maximised around 1983.

3.5. STEP FIVE - FORECASTING AND FORECAST EVALUATION

If the univariate modelling procedure is being utilised for forecasting purposes then this step can also form an important part of the diagnostic checking. Using ARIMA models for forecasting is relatively straightforward.

For example, consider a non-seasonal (1,0,1) model. The estimated model is given by

(8) $$X_t = \phi_1 X_{t-1} + a_t + \theta_1 a_{t-1}$$

Then the forecast value one period ahead conditional on all information up to time, t, is simply given by

(9) $$X^F_{t+1|t} = \phi_1 X_t + \theta_1 a_t$$

as $E_t(a_{t+i})$ equals zero $\forall$ i>0.

Similarly,

(10) $X^F_{t+2|t} = \phi_1 X^F_{t+1|t}$, as $E_t(a_{t+1})$ and $E_t(a_{t+2})$ equal zero and replace $X^F_{t+1|t}$ with the value given in equation (9).

(11) $X^F_{t+3|t} = \phi_1 X^F_{t+2|t}$

and so on.

To assess the out-of-sample forecasting ability of the model it is advisable to retain some observations at the end of the sample period which are not used to estimate the model. One approach is to estimate the model recursively and forecast ahead a specific number of observations. For example, consider a time series with data available from 1976Q1 to 1998Q4 and we wish to forecast four steps ahead (i.e., 1999Q1-1999Q4). We could initially estimate the model over the period 1976Q1 to 1992Q4 and forecast four steps ahead. Then re-estimate the model over the period 1976Q1 to 1993Q1 and forecast four steps ahead. Repeat this process until the estimation period leaves no out of sample observations available for forecast evaluation (i.e., 1976Q1 to 1998Q4). Using the actual inflation data over the period 1992Q1 to 1998Q4, this allows us to calculate 24 one step ahead forecast errors, 23 two-step ahead forecast errors, ..., and 21 four-step ahead forecast errors. These can be used to calculate statistics such as mean error (ME), mean absolute error (MAE), root mean squared error (RMSE) and Theil's U.

Denoting the forecast error as $e_t = X_t - X^F_t$ (i.e., the difference between the realised value of the series and the forecast value), then

(12) ME = $\frac{1}{F}\sum_{i=1}^{F} e_t$

(13) MAE = $\frac{1}{F}\sum_{i=1}^{F} |e_t|$

(14) RMSE = $\sqrt{\frac{1}{F}\sum_{i=1}^{F} (e_t)^2}$

(15) Theil's U = $$\sqrt{\frac{1}{F}\sum_{i=1}^{F}\left(e_t\right)^2} \Big/ \sqrt{\frac{1}{F}\sum_{i=1}^{F}\left(e_t^N\right)^2} = RMSE \big/ RMSE^N$$

where,

F equals the number of out-of-sample observations retained for forecast evaluation allowing for the forecast step, and

N denotes the naive model of no change in the modelled series from the last available observation.

One indication that the model specification could be improved is if the ME for each of the five steps are either all positive or all negative. This would indicate that the model is either forecasting too low on average (if positive) or too high on average (if negative).

If the ME is of the same magnitude as the MAE this would also indicate that the model is forecasting consistently either too low (if the ME is positive) or too high (if the ME is negative).

The RMSE will always be at least as large as the MAE. They will only be equal if all errors are exactly the same. Theil's U statistic calculates the ratio of the RMSE of the chosen model to the RMSE of the 'naive' (i.e., assuming the value in the next period is the same as the value in this period - no change in the dependent variable) forecasting model.[15] Thus, a value of one for the Theil statistic indicates that, on average, the RMSE of the chosen model is the same as the 'naive' model. A Theil statistic in excess of one would lead one to reconsider the model as the simple 'naive' model performs better, on average. A Theil statistic less than one does not lead to automatic acceptance of the model, but does indicate that, on average, it performs better than the 'naive' model. The advantage of the Theil statistic is that it is

[15] The naive model for the one-step ahead forecast assumes inflation follows a random walk (i.e., $\Pi_t = \Pi_{t-1} + e_t$).

‘unitless’ as it compares the RMSE of the chosen model to that of the ‘naive’ forecast model. The ME, MAE and RMSE all vary depending on the dimension (or scale of measurement) of the dependent variable. The Theil statistic also provides a quick comparison with the ‘no change’ model and, as such, is a measure for one-step ahead forecasts of the additional forecasting information the model provides beyond a random walk model.

An additional test of the ARIMA model would be to compare its performance with competing models including alternative ARIMA specifications and multivariate models.

These tests should be carried out on the range of models carried over from the diagnostic checking phase. Should some models forecast significantly worse than others this may be an indication of parameter instability or unit root problems if some of the factors of the AR or MA polynomials are close to or greater than unity.

Table 6 below presents some forecast statistics for DLHICP series estimated over the period 1984Q1 to 1998Q4. These statistics were calculated by first estimating the model over the period 1984Q1 to 1992Q4 and forecasting four steps ahead. The model was then recursively estimated, stepping forward one quarter at a time, and again forecast four steps ahead. As the sample period reaches 1998Q1, obviously one cannot forecast four steps ahead. Hence the number of forecast observations available declines with each step. The main points to note from Table 6 are as follows:

- The RMSE varies between 0.43% and 0.38%.[16] This implies a 90 per cent confidence interval of approximately 1.4 per cent per quarter. Although this appears high relative to the mean of the series, it compares favourably with results reported by Cecchetti (1995) for the United States who calculated a 90 per cent confidence interval of approximately 1.3 per cent for one step ahead inflation forecasts by a commercial inflation forecaster. These results also compare favourably with Bayesian vector autoregression (BVAR) results for forecasting at

[16] The mean of the DLHICP series over the period 1993 Q1 - 1998 Q4 was 0.56 per cent.

short horizons the Irish HICP as reported by Kenny *et al* (1998). For longer horizons (i.e., in excess of 4 quarters) the BVAR model outperforms the ARIMA model.

- The mean errors are significantly lower than the mean absolute error. This implies that the forecasts are neither systematically over-forecasting or under-forecasting inflation.
- The sign of the mean error varies by step. Again this implies that the forecasts are neither systematically over-forecasting or under-forecasting inflation.
- The Theil statistics are consistently below unity, indicating that the selected model outperforms the simple naive model. However, note that a straight line model (i.e., (0,0,0)(0,0,0) model) also outperforms the naive model.

In fact the fitted model (0,0,0)(1,0,1) is little more than a straight line model with allowance for a degree of seasonality.[17] Given the relatively stable pattern of inflation over the period, it is perhaps unsurprising that this model outperforms more elaborate models. It does indicate that, over the period 1993Q1-1998Q4, it would be difficult to improve on a very simple model of inflation. Thus over a short horizon forecast statistics from multivariate models are unlikely to outperform in any significant way the forecast statistics presented here for a relatively simple ARIMA model.

TABLE 6 - FORECAST STATISTICS FOR (0,0,0)(1,0,1) MODEL OF DLHICP (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.02	0.32	0.43	0.62	24
2	0.03	0.32	0.43	0.57	23
3	0.01	0.30	0.42	0.56	22
4	-0.03	0.28	0.38	0.71	21
Avg. 1-4	0.01	0.30	0.42	0.62	

Tables 7 and 8 present forecast statistics using DLHICPC and DsDLHICPC over the same period. Examining the forecast statistics for DLHICPC first, the impact of seasonal sales is shown up in the Theil statistics for the second and fourth step forecasts, which are significantly higher than those for other steps. This is directly as a result of the volatile nature of the series, with increases in one period generally being followed by decreases in the next. Hence, outperforming the naive forecast at odd steps would be expected. However, at even steps (i.e., when Summer and Winter sales and there respective rebounds coincide) the naive model will perform relatively well. Also note that the RMSE of the forecasts is higher than for the overall HICP indicating the greater volatility of the series. The RMSE vary between 0.66% and 0.69% (compared to approximately 0.42% for the overall HICP). There is little difference between Tables 7 and 8. The RMSE and Theil statistics are broadly similar, perhaps, reflecting the degree of ambiguity present in the tests for stationarity of the DLHICPC series. The results are obtained for the DLHICPE (services) series are somewhat similar to those for the DLHICPC series. Based on these results perhaps consideration of Autoregressive Fractionally Integrated Moving Average (ARFIMA) models for modelling the DLHICPC and DLHICPE series might be justified.

TABLE 7 - FORECAST STATISTICS FOR (0,0,0)(1,0,0) MODEL OF DLHICPC (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	-0.11	0.52	0.69	0.39	24
2	-0.07	0.49	0.67	0.96	23
3	-0.15	0.48	0.66	0.38	22
4	-0.18	0.47	0.66	1.06	21
Avg. 1-4	-0.13	0.49	0.67	0.70	

[17] The estimated seasonal AR coefficient is 0.58 and the estimated seasonal MA coefficient is -0.32.

TABLE 8 - FORECAST STATISTICS FOR (0,0,0)(1,1,1) MODEL OF DLHICPC (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.17	0.52	0.66	0.37	24
2	0.20	0.53	0.67	0.95	23
3	0.14	0.49	0.62	0.36	22
4	0.10	0.46	0.59	0.95	21
Avg. 1-4	0.15	0.50	0.64	0.66	

The forecast results for the other series are presented in Appendix C. The series with the highest RMSE, at 1.9 per cent, is the unprocessed food (HICPA). This compares to a mean of 0.64 per cent for that series. Seasonally differencing the overall HICP does not improve the forecast statistics. The RMSE increases from an average of 0.42 per cent to 0.58 per cent. All of the ARIMA models chosen outperform the simple naive model. This is true even for the straight line fitted to the energy series (HICPE), indicating that a Theil statistic of less than unity does not necessarily imply that a strong model has been found.

In general for the HICP and its sub-component series considered in this paper, a relatively parsimonious ARIMA representation has been found to be optimal both for fitting the in-sample data and for maximising the out-of-sample forecasting performance. This is an indication of the relative stability of inflation during the period in question and the dominance of the seasonal influence.

4. OUTLINE OF SEMI-AUTOMATIC ARIMA MODELLING ALGORITHM[18]

This paper has outlined an approach to ARIMA modelling. The approach is as follows:

---

[18] For a fully automatic univariate modelling procedure see Gómez and Maravall (1998). Their program also allows for outlier detection, which is not available in the program developed for this paper. Nonetheless the forecast statistics for the HICP generated using the semi-automatic algorithm are broadly similar to those generated using the automatic program by Gómez and Maravall.

- First, plot data and transformations of the data (consider possibility of outliers and structural breaks);
- Second, test for stationarity using both formal and informal tests;
- Third, run automatic algorithm which does the following:[19]
  - Estimates all 64 ARIMA models encompassed by (3,0,3)(1,0,1) model;
  - Calculate information criterion and Q statistic for all 64 models;
  - Provided models have properly converged and contain no unit roots in the AR and MA polynomials, rank models according to BIC;
  - Select top 10 performing models and compute forecast statistics using out-of-sample data;
  - Select model which optimises forecast performance.

5. Concluding Comments

This paper has considered autoregressive integrated moving average (ARIMA) forecasting. ARIMA models are theoretically justified and can be surprisingly robust with respect to alternative (multivariate) modelling approaches. Indeed, Stockton and Glassman (1987, pg. 117) upon finding similar results for the United States commented that "it seems somewhat distressing that a simple ARIMA model of inflation should turn in such a respectable forecast performance relative to the theoretically based specifications."

A framework for ARIMA modelling is identified which includes the following steps: data collection and examination; determining the order of integration; model identification; diagnostic checking; and, forecast performance evaluation.

Two alternative approaches to model identification are considered: the traditional Box-Jenkins approach which can be highly subjective; and, the objective penalty function criterion. The approach which is considered to be most robust is to retain a

[19] A copy of the RATS procedure used for this algorithm is available on request from the authors.

range of models, which perform satisfactorily at the model identification and diagnostic checking phases, for use in forecast performance evaluation. This approach is open to criticism of 'model mining', but since forecast performance is the overriding objective, this approach can be justified.

A general rule of thumb for univariate forecasting is to test, test and test at all stages of the process. Tests should be carried out across different time periods. Parameter stability should be determined. If stationarity is in doubt estimate the model in first differences and with seasonal differencing.

A semi-automatic algorithm has been developed for fitting an ARMA model to stationary time series data. The advantage of this algorithm is that it uses the objective penalty function criteria to select the optimal ARMA model, thus removing the subjectivity associated with the traditional Box-Jenkins methodology. It also recognises the conflict between model goodness-of-fit and out-of-sample forecasting performance. The algorithm is not fully automatic as it is always vital that the time series analyst undertakes a rigorous check of the models for consistency over time and for error autocorrelation.

Although the forecasting results for the sample period 1993Q1-1998Q4 compare quite favourably with those derived from BVAR analysis, this does not mean that univariate modelling can supplant multivariate techniques. The period in question was one of relatively stable inflation. ARIMA models may not perform as well with more volatile series. Furthermore, ARIMA models are 'backward looking' and are generally poor at forecasting turning points. Also well-specified multivariate models generally perform better than ARIMA models over longer time horizons.

One possible way to improve the forecasting performance is to attempt to fit an ARIMA model to a 'noiseless' version of the HICP series. This would involve applying statistical techniques, following Bryan and Cecchetti (1993), to remove 'noisy' random fluctuations from the measured inflation rate. This has the advantage that with the noise removed there would be less likelihood of parameter estimates being distorted due to outliers or other forms of noise. Preliminary analysis of a constructed ‘noiseless’ series, indicates that the optimal ARIMA model does indeed outperform the ARIMA model fitted to the noisy series but only marginally so.

APPENDIX A - A BRIEF OVERVIEW OF ARIMA MODELS

A general class of univariate models is the Autoregressive Integrated Moving Average (ARIMA) model. An ARIMA model represents current values of a time series in terms of past values of itself (the autoregressive component) and past values of the error term (the moving average terms). The integrated component refers to the number of times a series must be differentiated to induce stationarity.[20]

A.1. AR MODELS

A pure AR(p) process may be represented as follows, where X_t is modelled as lagged values of itself plus a 'white noise' error term.

(A1) $$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + a_t = \sum_{i=1}^{p} \phi_i X_{t-i} + a_t$$

This may be written alternatively as,

(A2) $$\phi(B) X_t = a_t$$

where, $\phi(B)$ is a p-order polynomial in the backshift operator (i.e., $1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$), and B is the backshift operator, such that $B^0 X_t = X_t$, $B^1 X_t = X_{t-1}$, $B^2 X_t = X_{t-2}$, ...

A useful way of gaining insight into univariate processes is to consider their autocorrelation and partial autocorrelation functions (ACF and PACF).

[20] In what follows, the modelled variable (X_t) is assumed to have been differenced sufficiently to achieve stationarity. For our purposes we define stationarity as second-order or weak stationarity. This requires that $E(X_t) = \mu \; \forall \; t$ and that $V(X_t) = \sigma^2 \; \forall \; t$. Furthermore, $COV(X_t, X_{t+k})$ depends only on the lag length, k, and not on time, t.

The ACF measures the ratio of the covariance between observations k lags apart and the geometric average of the variance of observations (i.e., the variance of the process if the process is stationary, as $V(X_t) = V(X_{t-k})$).

(A3) $$\rho_k = \frac{Cov(X_t, X_{t-k})}{\left[V(X_t) \bullet V(X_{t-k})\right]^{1/2}} = \frac{\gamma_k}{\gamma_0}$$

The sample autocorrelation function (SACF) may be calculated as follows[21];

(A4) $$r_k = \frac{\sum_{t=k+1}^{n}(X_t - \overline{X})(X_{t-k} - \overline{X})}{\sum_{t=1}^{n}(X_t - \overline{X})^2}$$

However, some of the observed autocorrelation between X_t and X_{t-k} could be due to both being correlated with intervening lags. The PACF seeks to measure the autocorrelation between X_t and X_{t-k} correcting for the correlation with intervening lags. For example, consider an AR(1) process of the form $X_t = 0.8\ X_{t-1} + a_t$. The first order autocorrelation coefficient is 0.8. The autocorrelation coefficient for the second lag is 0.64 (i.e., 0.8 * 0.8), although the partial autocorrelation coefficient for the second lag is zero, as the process is an AR(1) process. In other words the autocorrelation between observations two lags apart is due only to the correlation between observations one lag apart which feeds through into the second lag. As the lag length increases the autocorrelation coefficient declines (at lag length k the autocorrelation coefficient is $(0.8)^k$).

The PACF is calculated as the partial regression coefficient, $\phi_{kk,}$ in the kth order autoregression

(A5) $$X_t = \phi_{k1} X_{t-1} + \phi_{k2} X_{t-2} + \ldots + \phi_{kk} X_{t-k} + a_t$$

Thus, for an AR(p) process, $\phi_{kk} = 0 \; \forall \; k > p$.

Some general properties of the ACF and PACF for AR processes can be observed by considering a simple AR(1) process.

(A6) $X_t = \phi_1 X_{t-1} + a_t$

Note that the AR(1) model can be written as an infinite length MA process, providing $\phi_1 <$ unity. Denote the AR(1) series as,

(A7) $(1 - \phi_1 B) X_t = a_t$, where B is the backshift operator as before, which gives

(A8) $X_t = (1 - \phi_1 B)^{-1} a_t$ which upon expansion and providing $\phi_1 <$ unity yields

(A9) $X_t = a_t + \phi_1 a_t + \phi_1^2 a_{t-2} + \phi_1^3 a_{t-3} + \ldots$

This result holds more generally so that any finite order stationary AR process may be expressed as an infinite order MA process. This duality between AR and MA processes is an important property which can often be exploited when attempting to identify ARMA models.

For the AR(1) process the value of the ACF at lag k is given by ϕ_1^k. The value of the autoregressive coefficient can yield some insight into the underlying data generating process. For example, higher values of ϕ_1 indicate a higher degree of persistence in the series. A negative autoregressive component indicates a process which oscillates around its mean value.

For more general AR(p) models, the behaviour of the process is determined by the solution to the p-order polynomial ($1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$), given by

---

[21] Note, for large n, r_k is approximately normally distributed, so that a value in excess of $2(n)^{-1/2}$ can be regarded as significantly different from zero.

(A10) $\phi(B) = (1 - g_1 B)(1 - g_2 B)\ldots(1 - g_p B) = 0$

For the process to be stationary it is a necessary and sufficient condition for the roots of the p-order polynomial to lie outside the unit circle, i.e., $\left| \frac{1}{g_i} \right|$ > unity ∀ i = 1,...,p).

A.2. MA MODELS

A MA(q) process may be represented as follows, where X_t is modelled as the weighted average of a 'white noise' series.

(A11) $X_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + .. + \theta_q a_{t-q} = \sum_{j=0}^{q} \theta_j a_{t-j}$

or alternatively,

(A12) $X_t = \theta(B) a_t$

where,

a_t is a 'white noise' series,

$\theta(B)$ is a q-order polynomial in the backshift operator (i.e., $1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q$), and

B is the backshift operator, $B^0 a_t = a_t$, $B^1 a_t = a_{t-1}$, $B^2 a_t = a_{t-2}$, ...

Note that the expected value of X_t equals zero. Furthermore, the autocorrelation between X_t and X_{t+k} equals zero for k greater than q. Thus the order of the MA process, q, indicates the 'memory' of the process. All MA processes are stationary, regardless of the coefficients of the model. However, to ensure invertibility of the model (i.e., that the finite order MA process can be written in terms of a stationary infinite order AR process) the roots of the MA polynomial must lie outside the unit circle. MA models can be particularly useful for representing some economic time series as they can handle random shocks such as strikes, weather patterns, etc..

A.3. ARMA MODELS

An ARMA(p,q) series may be represented as

(A13) $X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \ldots - \phi_p X_{t-p} = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + .. + \theta_q a_{t-q}$

or alternatively

(A14) $\sum_{i=0}^{p} \phi_i X_{t-i} = \sum_{j=0}^{q} \theta_j a_{t-j}$

where ϕ_0 and θ_0 =1.

or more compactly

(A15) $\phi(B) X_t = \theta(B) a_t$

Using mixed ARMA models can be useful as it should usually be possible to represent a time series satisfactorily using fewer parameters than might be required with a pure AR or pure MA model.

A.4. SEASONAL ARMA MODELS

Seasonal data may be also modelled. The number of seasonal AR and MA terms are usually denoted by P and Q respectively. Thus, a general seasonal ARMA model may be represented as,

(A16) $\phi(B) \Phi(B) X_t = \theta(B) \Theta(B) a_t$

where,

$\Phi(B) = 1 - \Phi_{1s} B^{1s} - \Phi_{2s} B^{2s} - .. - \Phi_{Ps} B^{Ps}$

$\Theta(B) = 1 + \Theta_{1s} B^{1s} + \Theta_{2s} B^{2s} + .. + \Theta_{Qs} B^{Qs}$

s = the seasonal span, hence for quarterly data s = 4 and for monthly data s = 12.

A.5. ARIMA MODELS

The integrated component of an ARIMA model represents the number of times a time series must be differenced to induce stationarity. A general notation for ARIMA models is ARIMA (p,d,q)(P,D,Q), where p denotes the number of autoregressive terms, q denotes the number of moving average terms and d denotes the number of times a series must be differenced to induce stationarity. P denotes the number of seasonal autoregressive components, Q denotes the number of seasonal moving average terms and D denotes the number of seasonal differences required to induce stationarity.

This may be written as

(A17) $\phi(B)\Phi(B)\nabla^{d}\nabla_{s}^{D}Y_{t}=\theta(B)\Theta(B)a_{t}$

where,

X_t = $\nabla^{d}\nabla_{s}^{D}Y_{t}$ is a stationary series, and

$\nabla^{d}=(1-B)^{d}$ represents the number of regular differences and $\nabla_{s}^{D}=\left(1-B^{s}\right)^{D}$ represents the number of seasonal differences required to induce stationarity in Y_t.

Two important properties of the parameters of ARMA models are worth repeating. First, for an ARMA process to be stationary it is required that the modulus of the roots of the p-order AR polynomial be greater than unity (i.e., $\left|1/g_j\right| \rangle 1, \; \forall$ j = 1,...,p). Second, for an ARMA model to be invertible (i.e., representable as a stationary infinite lag AR model) the roots of the q-order MA polynomial should also be greater than unity (i.e., $\left|1/g_j\right| \rangle 1, \forall$ j = 1,...,q).

A.5. THEORETICAL JUSTIFICATION FOR ARIMA MODELS[22]

In the subsections above no theoretical justification for the use of ARIMA models was given. However, ARIMA models may be theoretically justified by recourse to Wold's Decomposition. Wold's Decomposition implies that any purely stochastic zero mean, covariance-stationary process, X_t, admits the representation

$$\text{(A18)} \quad X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} + \kappa_t = \sum_{j=0}^{\infty} \psi_j B^j a_t + \kappa_t$$

where,

a_t is 'white noise' and represents the error made in forecasting X_t on the basis of a linear function of lagged X (i.e., removing the deterministic component),

$\sum_{j=0}^{\infty} \psi_j^2 < \infty$,

κ_t is uncorrelated with a_{t-j}, though κ_t can be predicted arbitrarily well from a linear function of past values of X and is referred to as the linearly deterministic component of X_t, and

B^j is the backshift operator.

In principle the Wold Decomposition requires us to fit an infinite number of parameters, ψ_j. However, the function $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ can be approximated to any degree of accuracy by a quotient of finite order polynomials

$$\text{(A19)} \quad \psi(B) \approx \frac{\theta(B)}{\phi(B)} = \frac{1 + \theta_1 B + \ldots + \theta_q B^q}{1 - \phi_1 B - \ldots - \phi_p B^p}$$

Hence,

$$\text{(A20)} \quad X_t \approx \frac{\theta(B)}{\phi(B)} a_t$$

---

[22] This section is drawn from Hamilton (pg. 108, 1994).

Appendix B - Additional Penalty Function Statistics

Table B1 - Top Five Models Based on BIC - DsDLHICP and DLHICPA (1984Q1 - 1998Q4)

Rank	DLHICP		DLHICPA	
	Model	**BIC**	**Model**	**BIC**
1	**(1,0,0) x (1,1,0)**	**-10.400**	**(0,0,2) x (1,0,1)**	**-7.693**
2	**(1,0,0) x (0,1,1)**	**-10.392**	**(3,0,0) x (1,0,1)**	**-7.675**
3	**(0,0,1) x (1,1,0)**	**-10.367**	**(0,0,2) x (0,0,0)**	**-7.669**
4	**(0,0,1) x (0,1,1)**	**-10.364**	**(2,0,0) x (1,0,1)**	**-7.661**
5	(0,0,1) x (0,1,0)	-10.356	**(0,0,0) x (1,0,1)**	**-7.652**

Table B2 - Top Five Models Based on BIC - DLHICPB and DLHICPD (1984Q1 - 1998Q4)

Rank	DLHICPB		DLHICPD	
	Model	**BIC**	**Model**	**BIC**
1	**(0,0,0) x (1,0,1)**	**-9.592**	**(0,0,0) x (0,0,0)**	**-7.833**
2	(0,0,1) x (1,0,1)	-9.529	(0,0,0) x (1,0,0)	-7.784
3	(1,0,0) x (1,0,1)	-9.529	(0,0,0) x (0,0,1)	-7.772
4	(0,0,2) x (1,0,1)	-9.504	(0,0,1) x (0,0,0)	-7.766
5	(2,0,0) x (1,0,1)	-9.501	(1,0,0) x (0,0,0)	-7.765

Table B3 - Top Five Models Based on BIC - DLHICPE and DsDLHICPE (1984Q1 - 1998Q4)

Rank	DLHICPE		DLHICPE	
	Model	**BIC**	**Model**	**BIC**
1	**(1,0,0) x (1,0,0)**	**-9.547**	**(3,0,0) x (1,1,1)**	**-9.322**
2	**(0,0,1) x (1,0,0)**	**-9.528**	(3,0,1) x (1,1,1)	-9.254
3	(1,0,0) x (1,0,1)	-9.483	(2,0,0) x (1,1,1)	-9.243
4	(1,0,1) x (1,0,0)	-9.481	(0,0,0) x (0,1,1)	-9.227
5	(2,0,0) x (1,0,0)	-9.479	(3,0,0) x (0,1,1)	-9.207

APPENDIX C - ADDITIONAL FORECAST STATISTICS

TABLE C1 - FORECAST STATISTICS FOR (0,0,1)(1,1,0) MODEL OF DLHICP (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.26	0.43	0.56	0.80	24
2	0.35	0.50	0.61	0.82	23
3	0.33	0.48	0.60	0.79	22
4	0.31	0.46	0.57	1.06	21
Avg. 1-4	0.31	0.47	0.58	0.87	

TABLE C2 - FORECAST STATISTICS FOR (0,0,2)(1,0,1) MODEL OF DLHICPA (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.36	1.43	1.77	0.62	24
2	0.45	1.45	1.81	0.48	23
3	0.30	1.56	2.00	0.79	22
4	0.20	1.51	1.97	0.96	21
Avg. 1-4	0.33	1.49	1.89	0.71	

TABLE C3 - FORECAST STATISTICS FOR (0,0,0)(1,0,1) MODEL OF DLHICPB (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.16	0.48	0.62	0.83	24
2	0.25	0.44	0.53	0.56	23
3	0.21	0.39	0.46	0.62	22
4	0.18	0.37	0.44	0.74	21
Avg. 1-4	0.20	0.42	0.51	0.69	

TABLE C4 - FORECAST STATISTICS FOR (0,0,0)(0,0,0) MODEL OF DLHICPD (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.06	0.77	1.09	0.67	24
2	0.08	0.78	1.11	0.73	23
3	0.03	0.76	1.11	0.76	22
4	0.04	0.80	1.14	0.76	21
Avg. 1-4	0.05	0.78	1.11	0.73	

TABLE C5 - FORECAST STATISTICS FOR (1,0,0)(1,0,0) MODEL OF DLHICPE (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	-0.14	0.49	0.66	0.40	24
2	-0.07	0.58	0.70	0.70	23
3	-0.11	0.57	0.70	0.41	22
4	-0.12	0.57	0.71	0.83	21
Avg. 1-4	-0.11	0.55	0.69	0.58	

TABLE C6 - FORECAST STATISTICS FOR (1,0,0)(1,1,1) MODEL OF DLHICPE (1993Q1 - 1998Q4)

Step	Mean Error	Mean Abs. Error	RMS Error	Theil U	N.Obs
1	0.32	0.68	0.86	0.52	24
2	0.28	0.67	0.85	0.85	23
3	0.31	0.69	0.87	0.51	22
4	0.31	0.70	0.88	1.02	21
Avg. 1-4	0.30	0.68	0.87	0.72	

REFERENCES

Akaike, H., 1974. "A New Look at Statistical Model Identification", *IEEE Transactions on Automatic Control*, AC-19, pp. 716-723.

Anderson, O., 1976. *Time Series Analysis and Forecasting*, London: Butterworths

Banerjee, A., J. Dolado, J. Galbraith and D. Hendry, 1993. *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*, Advanced Texts in Econometrics, Oxford University Press: Oxford.

Beguin, J-M., C. Gourieroux and A. Monfort, 1980. "Identification of a Mixed Autoregressive Moving Average Process: The Corner Method", in O.D. Anderson (ed.) *Time Series*, North-Holland: Amsterdam.

Box, G. and G. Jenkins, 1976. *Time Series Analysis: Forecasting and Control*, Holden Day: San Francisco.

Bryan, M., and S. Cecchetti, 1993. "Measuring Core Inflation", *NBER Working Paper Series* No. 4303, March.

Cecchetti, S., 1995. "Inflation Indicators and Inflation Policy", in B. Bernanke and J. Rotemberg (eds.), *NBER Macroeconomic Annual 1995*, MIT Press: London.

Chatfield, C., 1979. "Inverse Autocorrelations", *Journal of the Royal Statistical Society, Series A*, Vol. 142, pp. 363-377.

Dotsey, M. and P. Ireland, 1996. "The Welfare Cost of Inflation in General Equilibrium", *Journal of Monetary Economics*, Vol. 37, pp. 29-47.

European Central Bank (ECB), 1998. "A Stability-Oriented Monetary Policy Strategy for the ESCB", *Press Release*, 13 October.

Feldstein, M., 1996. "The costs and benefits of going from low inflation to price stability", *NBER Working Paper* No. 5469, February.

Frain, J., 1995. "Econometrics and Truth", *Central Bank of Ireland Technical Paper 2/RT/95*.

Gómez, V. and A. Maravall, 1998. "Automatic Modelling Methods for Univariate Series", *Banco de España Working Paper* No. 9808.

Gray, H., G. Kelley and D. McIntire, 1978. "A New Approach to ARMA Modelling", *Communications in Statistics*, B7(1), pp. 1-77.

Hamilton, J., 1994. *Time Series Analysis*, Princeton University Press: Princeton.

Hannan, E., 1980. “The Estimation of the Order of ARMA Process”, Annals of Statistics, Vol. 8, pp. 1071-1081.

Harris, R., 1995. *Using Cointegration Analysis in Econometric Modelling*, Prentice Hall: London.

Kenny, G., A. Meyler and T. Quinn, 1998. “Bayesian VAR Models for Forecasting Irish Inflation”, *Central Bank of Ireland Technical Paper 4/RT/98*.

Litterman, R., 1986. “Forecasting with Bayesian Vector Autoregressions - Five Years of Experience”, *Journal of Business and Economic Statistics*, January, No. 1, Vol. 4, pp. 25-38.

Ljung, G. and G. Box, 1978. “On a Measure of Lack of Fit in Time Series Models”, *Biometrika*, Vol. 66, pp. 67-72.

Meyler, A., G. Kenny and T. Quinn, 1998. “A Note on the Construction of an Historical (November 1975 - May 1998) HICP Series for Ireland”, Central Bank of Ireland Research Department Memorandum, 1/RDM/98.

Mills, T., 1993. *The Econometric Modelling of Financial Time Series*, Cambridge University Press: Cambridge.

Mills, T., 1990. *Time Series Techniques for Economists*, Cambridge University Press: Cambridge.

Perron, P., 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis", *Econometrica*, Vol. 57, No 6, pp. 1361-1401.

Poskitt, D. and A. Tremayne, 1987. “Determining a Portfolio of Linear Time Series Models”, *Biometrika*, 74, pp. 125-137.

Schwarz, G., 1978. “Estimating the Dimension of a Model”, *Annal of Statistics*, Vol. 6, pp. 461-464.

Stockton, D., and J. Glassman, 1987. “An Evaluation of the Forecast Performance of Alternative Models of Inflation”, *Review of Economics and Statistics*, Vol. 69, No. 1, February, pp. 108-117.

Tsay, R. and G. Tiao, 1985. “Use of Canonical Analysis in Time Series Model Identification”, *Biometrika*, Vol. 72, No. 2, pp. 299-315.