UNIVERSIDAD CARLOS III DE MADRID

working papers

PROPERTIES OF TWO U.S. INFLATION MEASURES (1985-2005) [*]

Eva Vicente Martínez [1]

Abstract

Analyses are presented of 84 quarterly observations 1/85-4/05 on two U.S. index numbers of nominal prices often employed to measure inflation. Analyses are designed to answer two key questions of interest to macroeconomists. Is inflation stationary (I(0)) or stochastically non-stationary (I(1))? If it is I(1), is it scalar or multivariate? Both measures of inflation are found clearly to be I(1) and, for these measures, inflation is found clearly to be scalar. The paper also illustrates univariate analysis procedures (and report standards) considered to be more effective and convincing than those found in the existing literature on inflation measures.

Keywords: Inflation, Integration Order, Cointegration, United States Economy, Data-based Time Series Analysis.

[1] Departamento de Estadística, Universidad Carlos III de Madrid
Tel.: 91-6249362; e-mail address: evicente@est-econ.uc3m.es

1 Introduction

Time series analyses are presented of the 84 quarterly observations 1/85-4/05 on two U.S. index numbers of nominal prices often employed to measure inflation. These analyses are designed to answer two key questions on recent U.S. inflation experience. First, is inflation stationary $\left(I\left(0\right)\right)$ or is it stochastically non-stationary $\left(I\left(1\right)\right)$? Second, if it is $I\left(1\right)$, is inflation scalar or multivariate? Both questions are of interest to macroeconomists, in their own right and for the design of research on the relationships between inflation and other variables. Both measures of inflation analyzed are found very clearly to follow $I\left(1\right)$ processes and, as far as these two measures are concerned, inflation is found very clearly to be scalar.

Besides these contributions to the study of U.S. inflation, the paper contributes an illustration of a set of practical univariate time-series analysis procedures (and standards for reporting), considered by the author to be substantially more effective and convincing than those found in the existing applied time-series econometrics literature on measures of inflation and the general level of nominal prices.

One of the variables studied is the implicit deflator for Gross Domestic Product (GDP), divided by 100, designated P. It is the ratio of nominal GDP divided by real GDP, official data for which were downloaded from the web site of the Bureau of Economic Analysis (U.S. Commerce Dept.). This variable is seasonally adjusted, since both numerator and denominator are officially seasonally adjusted. The second variable is the Consumer Price Index for All Urban Consumers: All Items, originally monthly and not seasonally adjusted, numbers for which were downloaded from the Federal Reserve Economic Data web site, though the series is originally produced by the Bureau of Labor Statistics (U.S. Dept. of Labor). This monthly series is aggregated to quarterly by taking the simple geometric mean of monthly values for each quarter, and is designated PC. A detailed Statistical Data

Appendix is available on request from the author.

Each of P and PC is an index number of nominal prices. Each implies a measure of inflation, $\nabla \ln P$ and $\nabla \ln PC$ respectively, where "ln" represents the natural logarithm, and $\nabla \equiv 1 - B$ is the difference operator, B the back-shift operator, such that $BY_t \equiv Y_{t-1}$ and $\nabla Y_t \equiv Y_t - Y_{t-1}$ for any time series Y at time t.

Inflation is a long-run concept, though both $\nabla \ln P$ and $\nabla \ln PC$ are measurements, that may or may not coincide with the appropriate long-run concept, depending on transient components. However, the long-run statistical properties of $\nabla \ln P$ and $\nabla \ln PC$ necessarily coincide with those of the corresponding inflation concepts. The first question formulated above can thus be restated as: Is $\nabla \ln P \sim I(d)$ for $d = 0$ or for $d = 1$? (the same question is formulated for $\nabla \ln PC$). Given that the data analyses clearly reveal that $d = 1$ for both inflation measures, the second question is framed as: Is $\nabla \ln P - \nabla \ln PC \sim I(d)$ for $d = 0$ or for $d = 1$? The data analysis of $\nabla \ln (P/PC) \equiv \nabla \ln P - \nabla \ln PC$ reveals that $d = 0$, which means that a cointegration relation, with coefficients $(1, -1)$, is found. That is, the stochastically non-stationary components of the two inflation measures are seen to be common, so that inflation, as measured by these two alternative statistics, is scalar, not bivariate.

It is remarkable that there is still apparently considerable debate among economists about the order of integration of inflation in the U.S., since the rate of inflation is one of the most frequently analyzed macroeconomic variables and the U.S. economy is the most researched economy in the world. Nevertheless, prominent research economists do reveal a surprisingly wide variety of beliefs on this matter.

A common assumption in theoretical work in macroeconomics is that the rate of inflation follows an $I(0)$ process; i.e. the effects of innovations in it are transitory. For

example, Clarida *et al.* (1999) appear to make this assumption, since they propose a model in which the inflation rate is expressed as deviations from its long-run level; see p. 1665. Such a long-run level does not exist unless inflation follows an $I(0)$ process. Furthermore, there are empirical studies, such as Clarida *et al.* (2000), in which it is explicitly assumed, without data analysis to justify it, that inflation follows an $I(0)$ process. The empirical results of the present paper strongly indicate that the $I(0)$ specification for inflation in the contemporary U.S. economy is false, which calls into question all conclusions of either of the cited papers, as well as of many others not cited that use the same erroneous specification.

Nelson and Plosser (1982) had a major impact on the applied econometrics literature and originated a debate, still apparently open today, about the order of integration of macroeconomic time series, including the rate of inflation. Many authors analyzed the same data as if it were somehow paradigmatic, but with widely differing results. That data was very old even at the time, running from the late 19th century until 1970. It is of questionable quality, at least for the years before 1929, when the official statistical system was much more primitive. This data is also characterized by massive anomalous incidents associated with the Two World Wars, the Great Crash of 1929 and the Great Depression, and it was probably affected by price controls at times. At least one author, Perron (1989), appears to find evidence in this data that the general level of nominal prices follows an $I(0)$ process (around a linear deterministic trend), though this author appears later to have changed opinion. Several authors, e.g. Andreou and Spanos (2003), Nelson and Plosser (1982), Perron (1997), Rose (1988) and Zivot and Andrews (1992), appear to find evidence in this data that the general level of nominal prices follows an $I(1)$ process, which implies that inflation follows an $I(0)$ process. And there is at least one author, Rudebusch (1992), who appears to find evidence in this same data to the effect that one cannot discriminate between the hypotheses

$I(0)$ and $I(1)$ for the general level of nominal prices. It is noteworthy that not one of the authors mentioned considered the possibility that the general level of nominal prices follows an $I(2)$ process, which implies that inflation follows an $I(1)$ process, the clear finding of this paper for recent data.

A few papers ask whether the inflation rate follows an $I(0)$ or an $I(1)$ process, rather than posing that question for the general level of nominal prices. Both Culver and Papell (1997) and Lee and Wu (2001) analyze monthly data on the Consumer Price Index (CPI) for 13 countries, for periods beginning in 2/57 and ending in 9/94 and 4/99 respectively. Their conclusions coincide. Both find in univariate analyses for the U.S. that the inflation rate follows an $I(1)$ process, though they also present results from panel analyses indicating $I(0)$. Ng and Perron (2001) is a theoretical paper and only presents empirical results, for the GNP deflators of the G7 countries, to illustrate how easy it is to apply their new statistical methods. They find $d = 1$ for the inflation rate calculated with the quarterly U.S. Gross National Product (GNP) deflator for 2/62-2/97, but they make no claim to relevance for their empirical results.

Henry and Shields (2004) treat quarterly CPI data for 1/60-4/01. This paper is mainly interested in the possible non-linearity of the processes underlying the data, but does not find non-linearity for the U.S.: it concludes that U.S. inflation follows an $I(1)$ process.

There are authors, primarily interested in matters pertaining to the dependence of nominal interest rates on inflation, who have published empirical results on the order of integration of different measures of inflation. Rose (1988) finds several inflation measures to follow $I(0)$ processes, though each of Evans and Lewis (1995), Crowder and Hoffman (1996) and Crowder and Wohar (1999) finds inflation to follow an $I(1)$ process.

The problem with all four papers is that the evidence offered on the order of integration of inflation is extremely slight, probably because this is not the central issue they are considering. These articles will thus be ignored in the rest of this paper.

The present author, in Vicente (2005), analyzed both annual and quarterly data for both $\ln P$ and $\ln PC$, in the period 1959-1999 and in various sub-samples, finding very clearly that both inflation rates follow $I(1)$ processes, both in the overall sample and in the sub-samples. Clear changes in structure were detected between the period after that of the Federal Reserve's so-called "New Operating Procedures", which ended roughly with 3/82, and both the period before those procedures were begun (3/79), and the period when the procedures were in force. The main change detected is a large reduction in variance in both variables after 3/82. The present paper is an extension of that work, considering a period well after the end of the New Operating Procedures.

As far as the present author is aware, only the doctoral theses by Valbuena (2002) and Vicente (2005) have previously posed the second question, of whether inflation is scalar or multidimensional. Valbuena analyzes the implicit deflators for each of the thirteen aggregate components of expenditure and production, and for GDP, with annual data for the Spanish economy in 1964-1996, finding that these series follow $I(2)$ processes with a very few exceptions. She also analyzes the ratios of each component deflator to the GDP deflator and finds that these follow $I(1)$ processes, thus finding that inflation is scalar in that case also. In Vicente (2005) annual and quarterly data for $\ln(P/PC)$ are analyzed for the sample 1959-1999 and various sub-samples, and $\ln(P/PC)$ is shown to follow an $I(1)$ process which, together with the finding that $\ln P$ and $\ln PC$ follow $I(2)$ processes, implies that inflation is scalar.

The contents of the remaining sections of this paper are as follows. Section 2 presents

the time-series analysis tools used in the present work and compares them critically with those characterizing the other empirical studies cited. Sections 3, 4 and 5 present reports of the univariate analyses respectively of $\ln P$, $\ln PC$ and $\ln(P/PC)$. Section 6 concludes.

2 Methods of time series analysis

The time-series analysis methods employed in this paper are an extended version of those introduced in the 1st edition of Box, Jenkins and Reinsel (1994) some 36 years ago. This section is meant to define them.

Also in this section, the tools and methods used in other empirical studies of measures of the U.S. general level of nominal prices or rate of inflation, are analyzed critically with reference to the tools and methods employed in this paper. The studies considered are: Andreou and Spanos (2003), Culver and Papell (1997), Henry and Shields (2004), Lee and Wu (2001), Nelson and Plosser (1982), Ng and Perron (2001), Perron (1989), Perron (1997), Rudebusch (1992) and Zivot and Andrews (1992). They are representative of the empirical literature related to this research, recently published or frequently cited, and published in prestigious professional journals. No comments on the tools and methods employed by Valbuena (2002) are made, because they are mostly the same as those of this paper.

The section is divided into subsections, 2.1 on representation and 2.2 on model-building.

2.1 Representation

By representation is meant the general mathematical form of the parametric models within which one chooses a member to describe a particular time series.

Let Y_t stand for an arbitrary time series. In this paper the particular cases are P_t, PC_t and P_t/PC_t, as defined in section 1 and treated respectively in sections 3, 4 and 5.

The natural-base logarithm is used for each of the variables of this paper. Thus the

form $z_t = \ln Y_t$ is considered in all three cases. This specification is well-justified both by theoretical and empirical reasons and is common in the empirical literature on index numbers of nominal prices. If chosen for P_t and PC_t, it is implied for P_t/PC_t.

The variable z_t is represented as the sum of a purely deterministic component (ξ_t) and a purely stochastic component (N_t), $z_t = \xi_t + N_t$. The deterministic term (ξ_t) is the sum of deterministic trend, seasonal and intervention components.

The deterministic trend component, if any, takes the parametric form of a finite-order polynomial in the time index t. The deterministic seasonal component, in the case of quarterly data, takes the parametric form $\alpha_1 \cos(\pi t/2) + \beta_1 \sin(\pi t/2) + \alpha_2 (-1)^t$, the first two terms representing frequency $f = 1$, the last term representing frequency $f = 2$.

Each intervention term i takes the general form $v_i(B)\xi_t^i$, where

$$v_i(B) = \omega_{s_i}^i(B)\big/\delta_{r_i}^i(B) \text{ for } \omega_{s_i}^i(B) = \omega_0^i - \omega_1^i B - \cdots - \omega_{s_i}^i B^{s_i} \text{ and } \delta_{r_i}^i(B) = 1 - \delta_1^i B - \cdots - \delta_{r_i}^i B^{r_i},$$

where s_i and r_i are non-negative integers and the omegas and deltas are real parameters. The long-run gain of the linear transfer function $v_i(B)$ is defined as $g \equiv v_i(1)$. The ξ_t^i in this paper may take the form of a unit impulse at time t^*, ξ_t^{I,t^*}, with unit value at $t = t^*$ and zero values at all other times, or a unit step beginning at t^*, ξ_t^{S,t^*}, with unit values at all $t \geq t^*$ and zero values at all other times. This general representation was introduced by Box and Tiao (1975). The denominator operator in $v_i(B)$ is not necessary in any of the cases found in this paper.

In almost all models discussed in this paper, the purely stochastic component N_t is assumed to follow the gaussian ARIMA(p,d,q) form, $\phi_p(B)\nabla^d N_t = \theta_q(B)a_t$, where p, d and q are non-negative integers, $\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$, the AR(p) operator, is assumed to be

stationary, that is, all solutions to $\phi_p(B) = 0$ have modulus strictly greater than one,

$\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$, the MA(q) operator, is assumed to be invertible, that is, all

solutions to $\theta_q(B) = 0$ have modulus strictly greater than one, and a_t iid (independently and

identically distributed) $N(0, \sigma_a^2)$, where a_t is called the innovation at time t. The phis and

thetas are real parameters. It is assumed that the AR(p) and MA(q) operators do not have

common factors. It is said that a time series z_t follows a stochastic process integrated of

order d, $z_t \sim I(d)$, when the variable $\nabla^d N_t$ follows a stationary, invertible stochastic process

as defined here.

Note that the ARMA(p,q) model form is used in this paper, though most of the

existing empirical literature employs pure AR(p) forms. The ARMA form is preferred,

because: (1) it allows a more parsimonious representation than the pure AR form for many

time series that occur in practice, and it should never yield a less parsimonious form, and

(2) over-differencing implies the presence of non-invertible MA(1) factors, so that the use of

MA forms, together with AR, helps one detect over-differencing. All papers cited in the

second paragraph of this section ignore MA structure, except Culver and Papell (1997) and

Henry and Shields (2004). This is surprising, given that one of the main reasons for using

MA(1) structure is that it helps clarify the order of integration, the central objective of all of

these papers.

The time series P is officially seasonally adjusted but, as seen in section 3, contains a

deterministic seasonal mean. The time series PC is not officially seasonally adjusted, but the

data analysis of section 4 detects a deterministic seasonal mean in an over-fitting diagnostic

exercise. In both cases, it is worthwhile to check for the possibility that the seasonal mean is

stochastic rather than deterministic. This is done only in otherwise apparently statistically

adequate models. At $f = 1$, remove the cosine and sine terms of the deterministic

specification, add the homogeneously non-stationary operator $\left(1+B^2\right)$ to the left side of the

model and add an operator of the form $\left(1-\lambda_1 B^2\right)$ to the right side of the model, then estimate

the resulting model and check for the non-invertibility $\left(\lambda_1=-1\right)$ of this last MA operator. At

$f=2$ a similar procedure is employed. The $(-1)^t$ deterministic term is removed, the

operator $\left(1+B\right)$ is added to the left side, the MA factor added $\left(1-\lambda_2 B\right)$ is checked for

non-invertibility $\left(\lambda_2=-1\right)$. The relevant test statistics are discussed below. The initial ideas

for this approach to seasonal mean were presented in Gallego (1995) and Gallego and

Treadway (1996).

The equation $\phi_p\left(B\right)\nabla^d\left(z_t-\xi_t\right)=\theta_q\left(B\right)a_t$ characterizes the representation used in

this paper. Consider the representation typical of most of the econometric literature on unit

roots, $\phi_p\left(B\right)\nabla^d z_t=\xi_t+a_t$, where all symbols are defined as before and MA structure is not

allowed. The critical error in specification, common to this literature, is that the deterministic

component is added to the innovation a_t rather than being subtracted from the variable z_t.

This implies that the meaning for z_t of each deterministic term depends on the value of d and

the order and parameters of the AR operator.

Consider the case when the researcher includes a deterministic linear trend term, say

$\alpha+\beta t$, in the autoregression. Under $d=0$, z_t, which was presumably intended to be taken

as stationary around *this* deterministic linear trend term, is actually taken to be stationary

around a deterministic linear trend term of the form $\alpha/\phi_p\left(1\right)+\beta t/\phi_p\left(B\right)$, which is

parametrically very complex, depending not only on the parameters α and β, but also on all

of the AR(p) parameters. When $d=1$, however, it is ∇z_t that has this peculiar deterministic

linear trend component and the deterministic trend component in z_t is quadratic unless $\beta=0$

is imposed. This is why the Dickey Fuller (1979), DF, and Augmented Dickey Fuller (1981), ADF, null hypothesis is always composite: $d = 1$ and $\beta = 0$. Similar complexity and the need for composite hypotheses arise for intervention terms. None of this useless complexity and need for composite hypotheses is present for the representation used in this paper.

Furthermore, observe that the commonly used equation can be rewritten as $\nabla^d z_t = \xi_t / \phi_p(B) + a_t / \phi_p(B)$. This specification requires that the same AR operator is to operate on the stochastic innovation and on each and every deterministic term as well, no matter what the value of d may be. This amounts to the use of utterly arbitrary specification and parameter restrictions.

The assumption of this representation seriously undermines all studies using the DF or ADF statistics when even the simplest deterministic terms are included.

2.2 Model-building

The data analysis of this paper employs a conscious iterative process, beginning simple and gradually complicating the model to adequately describe the data in statistical terms, with as few parameters as possible. This process involves five stages: (1) initial specification, based on the data, (2) efficient estimation by Exact Unconditional Maximum Likelihood, (3) statistical diagnosis, based on the data, (4) reformulation, when the model is found to be inadequate at stage 3, and (5) use of the model, for hypothesis testing e.g., when the model is found to be adequate at stage 3.

In most of the existing applied time-series econometrics literature, including all of the papers cited at the beginning of this section, little attention seems to be paid to the data. The model-building activity is not a process at all, and certainly not a process revealed in publication, but is limited to an exercise in estimation, with model-selection reduced to the application of a few formal statistical hypothesis tests. This description is not entirely fair for the Andreou and Spanos (2003) paper, because these authors do employ a certain iterative

10

process in their attempt to achieve statistical adequacy, though it still seems that they pay insufficient attention to the data.

The scarcity of background information on their data published by many of the authors, gives an idea of the slight importance conceded to the data they analyze. Nelson and Plosser (1982) do not indicate the exact definition of almost half of the series they analyze. Later papers, Andreou and Spanos (2003), Perron (1989), Perron (1997), Rudebusch (1992) and Zivot and Andrews (1992), analyze the same data but do not clarify the definitions. Culver and Papell (1997) do not indicate whether the monthly data on PC that they analyze is seasonally adjusted or not. There are even authors, such as Henry and Shields (2004), who do not mention the source of their data.

Both initial-specification and statistical-diagnosis operations presented in this paper make use of a single basic set of tools: the data graph, the graph of the autocorrelation function (*acf*), the graph of the partial autocorrelation function (*pacf*) and a small set of elementary statistics. These tools are applied to the level, the first difference and perhaps the second difference of the time series under analysis, at the initial-specification stage, and are applied to the residuals, the estimated innovations, at the statistical-diagnosis stage.

Let w_t stand for a generic variable and assume there are n observations on it (it might be z_t, ∇z_t, $\nabla^2 z_t$ or the residuals, $\hat{a}_t$). In the case of a particular variable, its precise name is found at the top of the graphics module, which contains the data graph, *acf* and *pacf* graphs and elementary statistics for that variable. Data graphs are presented standardized, that is, the variable presented is $(w_t - \bar{w})/\sigma_w$, where $\bar{w}$ stands for the simple arithmetic mean and σ_w stands for the standard deviation of the data. Thus the zero on the vertical axis marks the mean and the marks, numbered or not, refer to units of the standard deviation. Below each data graph, the two elementary statistics $\bar{w}$ and σ_w are presented, and in parenthesis after

the value of the mean appears the standard deviation of the mean $\sigma_{\bar{w}} = \sigma_w / \sqrt{n}$.

The *acf* is calculated for $k = 1, \cdots, 15$ in this paper. The *acf* coefficient at lag k, r_k, is

calculated as $r_k = \left(1/n\sigma_w^2\right) \sum_{t=1}^{n-k} \left(w_t - \bar{w}\right)\left(w_{t+k} - \bar{w}\right)$. The *acf* values are graphed together with

broken lines at $\pm 2 / \sqrt{n}$, which serve as guides to relevance. The Ljung and Box (1978)

portmanteau statistic Q is presented with the appropriate degrees of freedom in parenthesis.

The *pacf* is calculated for the same set of lags as the *acf*, and is graphed with the same

broken lines at $\pm 2 / \sqrt{n}$. The first value of the *pacf* is equal to that of the *acf*. For $k > 1$, the

pacf value is calculated as the Least Squares regression coefficient of w_t on w_{t-k} in a

regression with constant term and other regressors $\{w_{t-1}, \cdots, w_{t-k+1}\}$.

The analysis reports of sections 3-5 illustrate the use of these tools. At the initial

specification stage, one begins by considering the level of z_t. When the series clearly trends

(only crosses the mean value roughly once in the sample), it may do so upwards, as occurs

with $\ln P$ and $\ln PC$, or downwards, as occurs with $\ln\left(P/PC\right)$. Then one concludes that the

series follows a non-stationary process, though one does not yet know if the non-stationarity

is deterministic or stochastic in nature. When the *acf* is seen to die out, as the order of lag

rises, only very slowly and either linearly or concave downward, this confirms the conclusion

of non-stationarity. In such cases, one proceeds to examine the case of ∇z_t. The only tools

used here are the data graph and the *acf*. One does not take a definitive decision as to the

order of integration at all at this point, but takes a first step in a process by which, later on,

such a decision may be rigorously obtained.

One continues this process until one obtains a data graph that appears to be

well-centered and an *acf* that appears to arise from a stationary process, that is, the latter dies

out rapidly, either with a clear cut-off point or being strongly damped, and has a

configuration that appears to reflect an ARMA(p,q) structure with low orders (p,q). The basic ARMA forms of *acf* (and *pacf*) can be found in Box, Jenkins and Reinsel (1994). When one arrives at such evidence, without finding a need for intervention analysis along the way, then one proceeds to efficiently estimate the resulting ARIMA(p,d,q) model.

In many economic time series, as in the cases of this paper, the level of the series shows clear trend and hence one will initially use at least one difference. It is also true that the second difference of most economic series is well-centered, i.e. shows strong affinity to a long-run mean value, usually zero. However, there are many intermediate cases, as occur with $\nabla \ln P$ and $\nabla \ln PC$ in this paper. In such intermediate cases, the data graph neither shows trend, crossing the mean line only roughly once in the sample, nor does it show a well-centered series, crossing the mean line very frequently. The series is said to wander. In such cases, one may choose to try to model the series as stationary and look for a non-stationary AR(1) factor or try to model the series with one more difference and look for a non-invertible MA(1) factor. In really doubtful cases, one is best advised to consider both options, hoping that they will lead to the same conclusion. Note that one will often have to use a mean parameter (constant in ξ_t) in the potentially under-differenced case, but not in the potentially over-differenced case.

Seasonal variation in the local mean of the data can appear at any level of differencing. It is seen in the data graph when the values at a certain season (quarter here) are always, or almost always, above (or below) the mean (the zero axis). It is seen in the *acf* graph by positive values that do not seem to die out at the annual lags; they need not be large. When such evidence is found, one may proceed to initially model the seasonal mean either as deterministic or stochastic, testing the specification against the alternative at a later stage. In this paper, the initial specification is deterministic and it is found to be adequate in all cases, once the stochastic alternative is considered. This choice is made here, because there should

be no sign of seasonal mean in the variable P, since it is officially seasonally adjusted, and because seasonal mean is found in PC only at the end of the iterative model-building process, in an over-fitting statistical diagnosis operation.

One often detects anomalous incidents in the data as one carries out the process previous to efficient estimation. If such incidents are very important, leading to probable distortions in the analysis tools, especially the *acf*, but also in the elementary statistics, then one may efficiently estimate a model with no ARMA structure, but with intervention terms representing the anomalies, before proceeding with ARMA specification. The evaluation of such distortions is not detailed here, but basically amounts to calculating and evaluating the contribution of the data anomaly to the statistic expected to be distorted. In other cases, where such anomalies appear to not be so important, one may carry out intervention analysis after efficiently estimating an ARMA model.

Intervention analysis is useful, not only for controlling the distortion and influence potential of anomalous data, but also for discovering the exact form of the anomaly. One should, once the form of the anomaly is clear, look for explanations for the anomaly in whatever outside sources one can find. Such information is called extraneous information, because it does not derive from the sample data under analysis. If, in the case of a given anomaly, one finds no such extraneous information, then the decision to use the intervention terms in a final model should be based on an evaluation of the influences of the anomaly, influences on the choice of the order of integration, and/or influences on the other parameters of the model. At the very least, one must be aware of influences that are present and avoid taking decisions that depend on ignoring (not modeling) certain anomalous data.

Graphical materials are used intensively in the research reported in this paper, starting at the initial-specification stage. In the papers cited at the beginning of this section, graphical materials are almost completely ignored. Only Zivot and Andrews (1992) and Henry and

Shields (2004) include data graphs of all series they analyze, and Zivot and Andrews (1992) only present data graphs of the levels of the series. It is surprising that, with the exception of Henry and Shields (2004), in none of the studies cited are data graphs and *acf/pacf* of both levels and first differences presented, when the studies themselves are intended to evaluate whether the series follow $I(1)$ processes or $I(0)$ processes. These studies systematically ignore the most elementary evidence.

All efficient estimation of models in the present paper is performed with a computer program (FDRVUS) originally designed, coded and debugged by J. Alberto Mauricio, tested extensively in that original form by a team of researchers at the Universidad Complutense de Madrid and modified later by David Guerrero to simplify use and to free it from dependence on non-free code. The C source code together with user's manual and technical manual is available from the author. The estimation criterion is Exact Unconditional Maximum Likelihood and the implementation combines procedures for evaluation of the exact unconditional maximum likelihood function for the univariate ARMA model by Melard (1984) with other procedures developed by Mauricio (1992, 1995). This criterion is to be preferred to the conditional maximum likelihood criterion, especially when sample size is small, because: (1) the conditional criterion is but an approximation to the unconditional criterion and (2) the conditional criterion regards the initial p observations, of the variable taken as stationary for the estimation, as fixed values, so that residual values are not generated for these early dates in the sample, leaving the researcher open to the presence of undetected data anomalies in those early observations, anomalies that may well influence any of the inferences drawn. All papers cited above use conditional efficient estimation methods.

Statistical diagnosis of a model begins with writing the model down with the numerical values estimated for the parameters in the factored form of the operators, estimated standard errors in parenthesis below parameters, estimated correlations between parameters

presented when sufficiently large (greater than .7 in absolute value), estimated damping factor and period with their corresponding approximate estimated standard errors presented below any factors of second degree with imaginary roots, and the estimated standard error of the innovations. One evaluates the possibility of irrelevant parameters by comparing their estimated values with their standard errors; Student t statistics can be employed. The same factor must not appear in both the AR and MA operators; factoring operators allows one to quickly check for such parameter redundancy. Over-parameterization may be indicated by high correlations between parameters. The presence of an AR(2) factor with imaginary roots may suggest an over-fitting experiment by adding an MA(1) factor to the model. All of these steps lead potentially to reformulations, but do not involve examination of the residuals.

For a model to appear to be statistically adequate, none of the above checks should suggest possible reformulations not yet tried, and furthermore the residuals must appear to comply with the specification a_t iid $N\left(0,\sigma_a^2\right)$. The residual data graph should show a well-centered series with not too many large residuals, not more than roughly 2.5% of the number above (below) the $+2\hat{\sigma}_a$ $\left(-2\hat{\sigma}_a\right)$ band, none very extreme in absolute value. The residual mean should not be much larger in absolute value than its standard deviation, though the residual mean can be very distorted by one or a few large residuals or an asymmetry of large residuals, so that contributions should be calculated to try to discount such possibilities. The residual *acf/pacf* should not indicate any further ARMA structure to include; the low-order lags are the most important. There is no reason to try to model isolated *acf* values at high lags. The Ljung and Box (1978) portmanteau Q statistic should not be much larger than the degrees of freedom (order of highest lag used for *acf* less the number of ARMA parameters estimated). If it is, that suggests misspecification, but, if it is not, that does not indicate the model is well specified, just that that particular test does not suggest misspecification. Note that the Q depends on the terms of the *acf* and the *acf* depends on the

16

time series shown in the data graph. Often a few extreme residuals will distort certain values of the *acf* at irrelevant lags and hence will distort the Q upwards as well. When one has used intervention analysis to evaluate the influence of such extreme values, has found that they do not influence either the differencing decision or the parameters of the model and has found no extraneous information indicating that the extreme values are due to contamination of some kind, then one should not leave the intervention terms in the model, but may well have a "dirty" residual *acf* and a Q statistic suggesting misspecification, even though one has no misspecification at all.

Diagnosis by over-fitting is a useful technique. If one suspects, for whatever reason, that the model might be improved in a certain direction, then one tries the reformulation. One should never add more than one or two parameters at a time in over-fitting.

In sharp contrast to what is done in the present paper, in the articles cited above no graphical tools for diagnosis are used. Even formal hypothesis tests of specification are not presented, with the exceptions of Andreou and Spanos (2003) and Henry and Shields (2004). For these reasons, the reader cannot determine whether these models are statistically adequate or not and, hence, cannot determine whether the conclusions based on hypothesis tests about the order of integration are reliable or not.

Furthermore, with the exception of Rudebusch (1992), the cited articles do not even present all of the parameter estimates with their estimated standard errors, nor the correlations between parameters. This makes it impossible for the reader to determine whether or not the models are over-parameterized, in which case the results of hypothesis tests on the order of integration are not reliable.

Reformulation, if needed, arises from the nature of the deficiencies detected in the diagnosis of the latest efficiently estimated model. It is easy to see how one reformulates when the problem is to remove redundant AR or MA factors. When the residual *acf/pacf*

reveal AR or MA structure not yet included, then this further structure is specified according to what is seen in the *acf/pacf* and is incorporated multiplicatively with the existing AR and MA factors. When the estimation situation is found to be ill-defined, that is, when large correlations are found between parameters, reformulation can take a number of forms, depending on how badly over-parameterized the model is. Of course, the presence of data anomalies often motivates reformulations, either before ARMA structure has been specified or after. In such cases it is important to take the data into consideration, and to take the ARMA structure already included into consideration, when interpreting the form of the anomaly in the level variable z_t. It is often easier to see the form in data graphs before ARMA structure has been used.

One may perform informal, or even formal, hypothesis tests of different kinds throughout the iterative model-building process, but the tests of the key hypotheses, such as, e.g., tests to determine the order of integration, are carried out only toward the end of the process. For a formal hypothesis test to be of real use, it must be employed under the appropriate conditions: (1) in a model that is parsimoniously parameterized, (2) that has been efficiently estimated, and (3) that has been found to be statistically adequate. It is not possible for the reader to know whether these conditions are satisfied or not for the papers cited above, because they do not present the information needed for evaluating conditions (1) and (3).

Furthermore, to take the results of a formal hypothesis test seriously, they must be clear-cut. To clearly reject a null hypothesis, one needs the test statistic to be much larger than the critical value at 95% confidence, to not reject one needs it to be much smaller than the critical value at 90%, say. There are, of course, cases that arise in practice, in which clear-cut results are not available, but one cannot rely heavily on them. In this paper, critical values at both 90% and 95% confidence are cited in parenthesis along with each test statistic.

The test for the non-stationarity of an AR(1) factor used in this paper is by Shin and Fuller (1998), and is called the SF test. The test for the non-invertibility of MA factors used in this paper is by Davis et al. (1995, 1996), and is called the DCD test. Both are based on unconditional likelihood ratios and appear to be superior both in size and in power to the alternative tests currently available for these hypotheses. The critical values used for the SF test are linear interpolations of those tabulated in Shin and Fuller (1998) for different sample sizes: 1.06 (90%) and 1.75 (95%). The critical values tabulated by Davis *et al*. (1995) for an $MA(1)$, also used for testing $(\lambda_2 = -1)$ in the $(1 - \lambda_2 B)$ MA factor used in quarterly seasonality work, are 1.00 (90%) and 1.94 (95%). The critical values for testing non-invertibility $(\lambda_1 = -1)$ in the $(1 - \lambda_1 B^2)$ MA factor used in quarterly seasonality work, are linear interpolations of those tabulated for an $MA(1)$ in Davis et al. (1995) and those tabulated for the $MA(1)_4$ in Davis *et al*. (1996): 1.07 (90%) and 2.02 (95%).

When extraneous information is found that convincingly explains a certain data anomaly as some kind of contamination or error in the data, then the intervention terms that adequately represent the data anomaly are left in the model.

When a data anomaly is not explained by extraneous information, then the empirical question arises of what influences it has. The intervention parameters associated with the data anomaly are evaluated for two kinds of influence: (1) influence on the inference about the order of integration and (2) influence on the rest of the estimated parameters of the model. Such evaluation is carried out with an efficiently estimated, parsimoniously parameterized and apparently statistically adequate model. The model is estimated with and without the intervention term. An intervention parameter is considered to influence the decision about the order of integration when its exclusion from the model changes the decision. An intervention parameter is considered to influence another estimated parameter when the latter

parameter estimate changes substantially when the intervention parameter is removed; this influence is considered to be substantial when the change in the other estimated parameter is by more than one estimated standard error, though this rule is not applied rigidly.

With the exceptions of Perron (1989) and Henry and Shields (2004), the articles cited above do not present data graphs to look for possible anomalous incidents in the data. In the papers by Nelson and Plosser (1982), Rudebusch (1992), Lee and Wu (2001) and Ng and Perron (2001), the presence of anomalous incidents in the data is not even mentioned, though the samples analyzed by Nelson and Plosser (1982) and Rudebusch (1992) include the Great Depression and two world wars, and the samples analyzed by Lee and Wu (2001) and Ng and Perron (2001) include the oil crises of 1973 and 1979 and the period in which the Federal Reserve applied the "New Operating Procedures", all of which definitely are associated with major data anomalies in inflation measures. One must wonder if these authors fail to mention any data anomalies, because they have not ever looked at the data graphs that would reveal them.

In none of the papers cited is there any mention of extraneous information, nor are the influences of ignored data anomalies on the test results evaluated. To ignore the presence of important anomalous incidents in the data of a time series can easily lead to utterly erroneous conclusions as to the degree of integration and even to the conclusion that the series follows a non-linear stochastic process when, in fact, the process is linear.

3 Univariate analysis of GDP deflator (P)

The data graph of $\ln P$ shows an upward trend with slowly varying slope. The coefficients of the *acf* die out only slowly and linearly.

If one were to attempt to represent this series as stationary around a deterministic trend, one would need to employ a deterministic trend component with linear or possibly higher-order polynomial terms. One would seek an initial AR form, but the order of such an

20

AR form is highly doubtful, because there are large, though isolated, *pacf* coefficients up to high orders.

lnP

$$\overline{\mathrm{w}}(\hat{\sigma}_{\mathrm{w}}) = -10.48 \,\% \,(1.52 \,\%)$$
$$\hat{\sigma}_{\mathrm{w}} = 13.93 \,\%$$
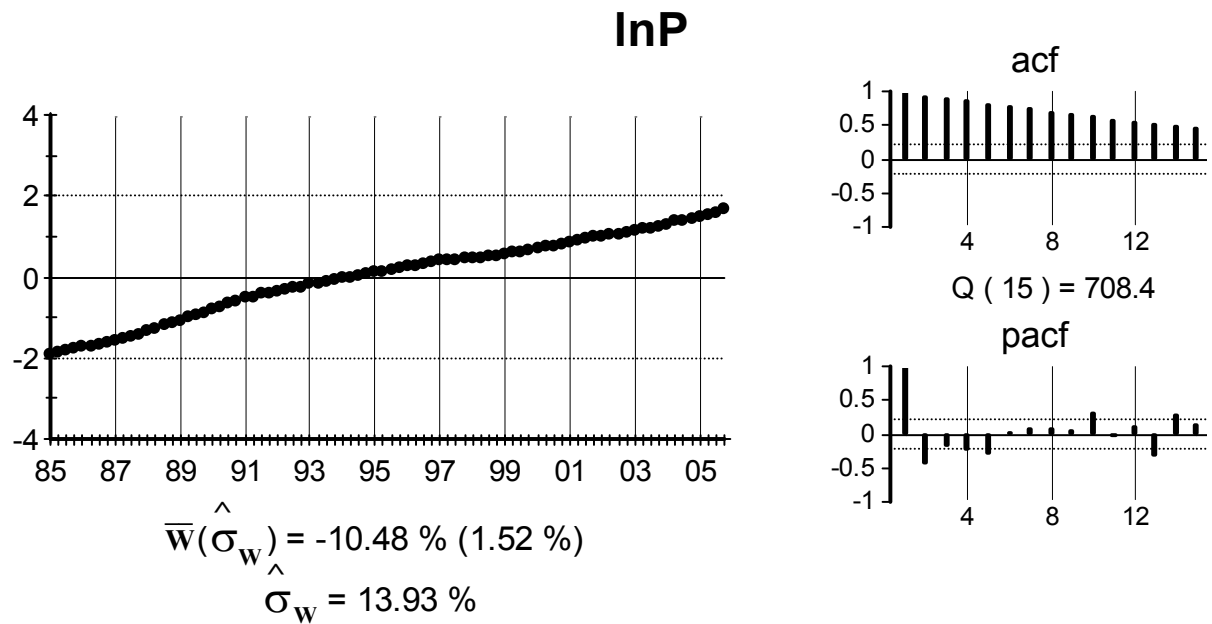
No definitive decision is yet taken on the order of differencing required. But the evidence suggests that $\ln P$ is probably stochastically non-stationary. It is thus wise to examine the $\nabla \ln P$ series.

∇ lnP

$$\overline{\mathrm{w}}(\hat{\sigma}_{\mathrm{w}}) = 0.60 \,\% \,(0.03 \,\%)$$
$$\hat{\sigma}_{\mathrm{w}} = 0.23 \,\%$$
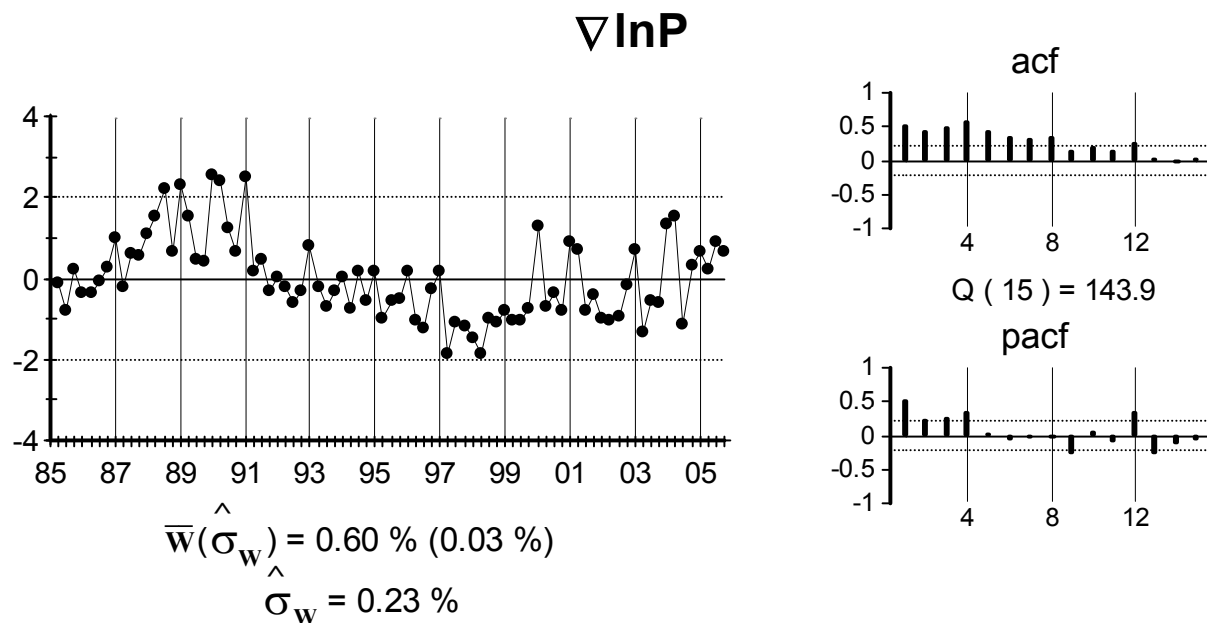
The data graph of $\nabla \ln P$ wanders very noticeably, showing little affinity for a long-run mean value. The *acf* is dominated by positive values that die out only very slowly.

This evidence suggests that even $\nabla \ln P$ probably follows a stochastically non-stationary

process. One should examine the $\nabla^2 \ln P$ series.

The data graph of $\nabla^2 \ln P$ appears to be well-centered, showing strong affinity for the

long-run mean value of zero. The *acf/pacf* suggest an MA(1) form. These features indicate

that $\nabla^2 \ln P$ follows a stationary stochastic process, though it may be non-invertible.

$\nabla^2 \ln P$

$$\overline{W}(\hat{\sigma}_W) = 0.00 \ \% \ (0.03 \ \%)$$
$$\hat{\sigma}_W = 0.23 \ \%$$
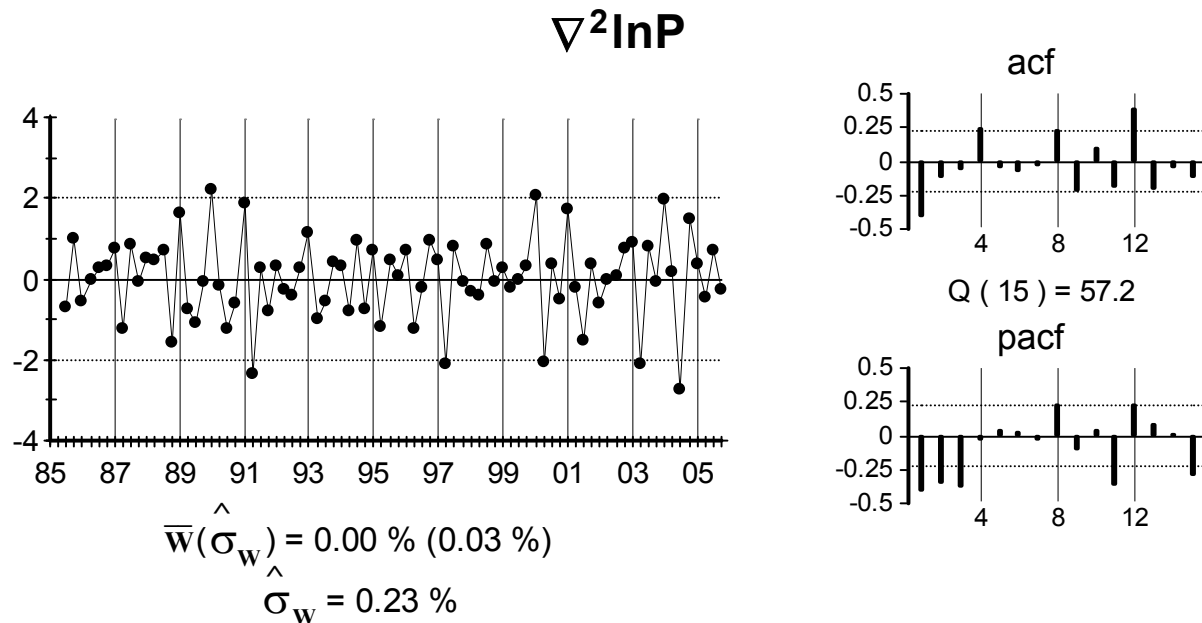
Note that there is evidence of a seasonal mean in $\nabla^2 \ln P$, though the time series is

officially seasonally adjusted. In the data graph, all first quarter values are positive, except

for small negative values in 1986, 1998 and 2002. In the *acf*, all three annual coefficients are

outstandingly positive. Note, also, that there are clear satellite values at lags 9, 11 and 13,

which reinforce the MA(1) specification and the impression of the presence of seasonality,

though there are no satellite values at lags 3, 5 or 7.

In the data graph, there are roughly nine values found at or beyond the $\pm 2\sigma$ bands,

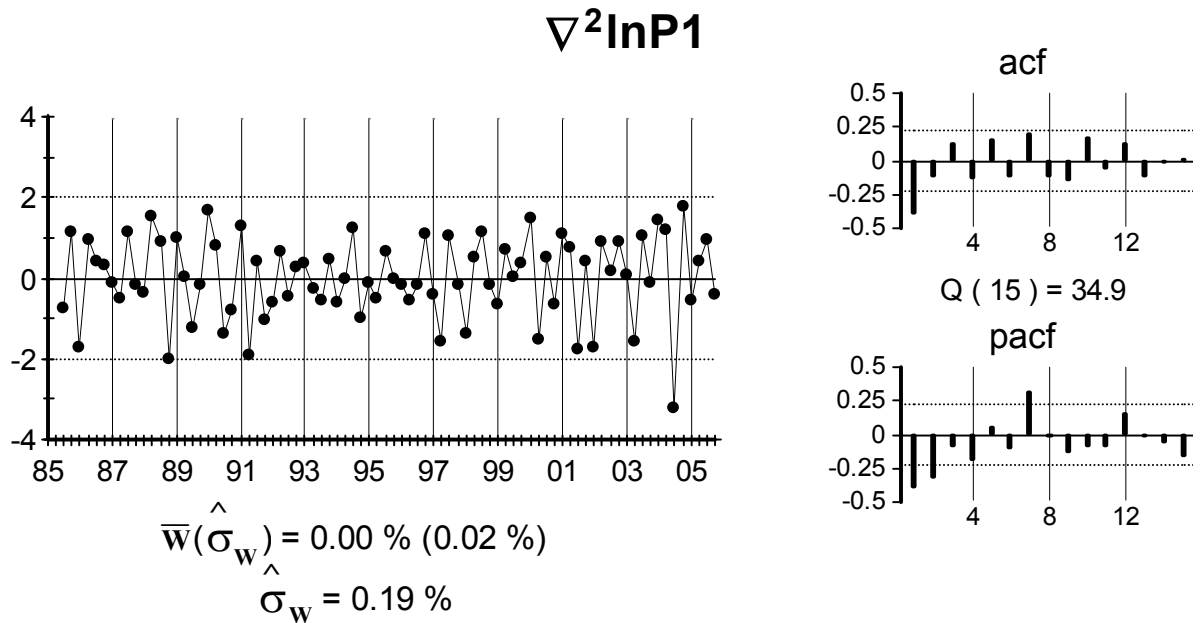not too many in a set of 82 values. None are very extreme.

The evidence of a seasonally varying mean in $\nabla^2 \ln P$, despite the fact that the

variable P is officially deseasonalized, requires attention. The approach taken here is to

initially assume that the seasonal mean is deterministic, complete a univariate analysis under

this assumption and then evaluate whether the seasonal mean is stochastic or deterministic.

To begin with, a model is estimated with deterministic seasonal mean under the assumption

that the remainder series $(\ln P1)$ follows a stationary, invertible process in the second

difference. Though not, of course, identical, the graphical and statistical materials for $\ln P1$

and $\nabla \ln P1$ are very similar to those for $\ln P$ and $\nabla \ln P$.

$$\ln P_t = -.00050\cos(\pi t/2) + .00048\sin(\pi t/2) - .00021(-1)^t + \ln P1_t$$
$$\quad\quad\quad (.00015) \quad\quad\quad\quad (.00015) \quad\quad\quad\quad (.00005)$$

(3.1)

$$\nabla^2 \ln P1_t = a_t$$

$$\hat{\sigma}_a = .19\%$$

$\nabla^2\ln P1$

acf

Q (15) = 34.9

pacf

$$\overline{W}(\hat{\sigma}_W) = 0.00\ \%\ (0.02\ \%)$$
$$\hat{\sigma}_W = 0.19\ \%$$

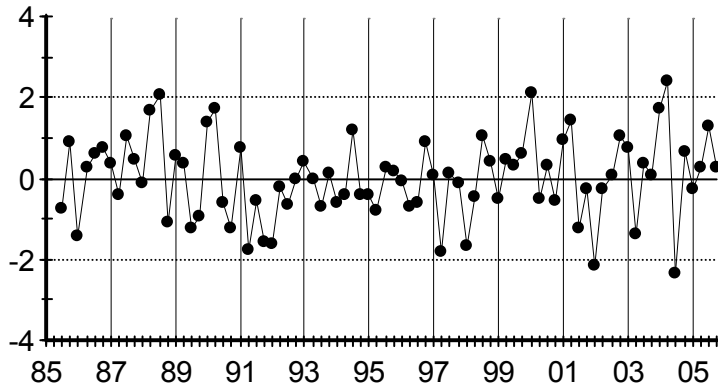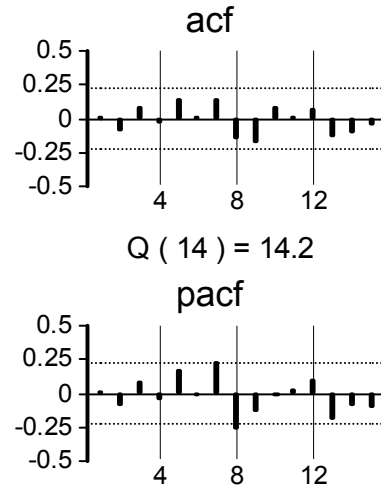It appears that the seasonal mean has been more or less adequately treated.

First-quarter values of $\nabla^2 \ln P1$ are not systematically of one sign. There is no evidence of

seasonal mean in the *acf*. The evidence of MA(1) structure with positive parameter is very

clear in the *acf/pacf*. There are three values on or beyond the $\pm 2\sigma$ bands and only the one at

3/04 seems possibly very extreme.

When the MA(1) structure is added, the estimated Model P2 follows.

AP2

$$\overline{w}(\hat{\sigma}_w) = 0.01\ \% \ (0.02\ \%)$$

$$\hat{\sigma}_w = 0.16\ \%$$

$$\ln P_t = -.00050\cos(\pi t/2) + .00048\sin(\pi t/2) - .00021(-1)^t + N_t$$
$$(.00015)(.00015)(.00007)$$

$$\nabla^2 N_t = (1 - .60B)\,AP2_t \tag{3.2}$$
$$(.08)$$

$$\hat{\sigma}_{AP2} = .16\%$$

Large Residuals

3/88	2.1σ
1/00	2.1σ
1/02	-2.2σ
2/04	2.4σ
3/04	-2.4σ

$$r_8 = -.15(.11)\ \{a_{1/00},\ a_{1/02} : -.06$$
$$r_9 = -.18(.11)\ \{a_{1/02},\ a_{2/04} : -.06$$

    This model appears adequate. The residuals seem well-centered with zero mean. There are but five residuals greater than the $\pm 2\sigma$ bands, not excessive for 82 values. No residuals are very extreme.

    The only anomalous incident that stands out a bit is a sequence of step-in-level effects in 1-3/04, but even this incident should not influence the parameter estimate for the MA(1), because the contributions of 1-2/04 and 2-3/04 are of opposite signs. Neither *acf* nor *pacf* nor Q statistic suggests misspecification. The only slightly large *acf* values, at lags 8 and 9, are seen to receive substantial contributions from just one interaction in each case. The estimated

parameter values for the deterministic seasonal components are exactly equal to those obtained without the MA(1) parameter.

Model P2 specifies $\ln P \sim I(2)$. To check for over-differencing, the hypothesis $H_0 : \theta_1 = 1$ is tested. The DCD statistic is 54.74, which is much larger than either of the critical values of (1.00, 1.94) for confidence levels (90%, 95%), thus rejecting the hypothesis. When an AR(1) factor is added, the parameter is not significantly different from zero and the DCD statistic is 11.26, confirming the rejection of the over-differencing hypothesis.

When three step-in-level terms in 1-3/04 are added to Model P2, the residuals in the neighborhood of these dates indicate that any anomaly present is adequately represented, and no parameter of the model varies by more than one standard deviation (as estimated in P2), indicating that, as expected, this incident is not influential in the parameter estimates. The DCD statistic for the test of $H_0 : \theta_1 = 1$ is found to be 65.0, to be compared with the same critical values given above, so that the over-differencing hypothesis is strongly rejected, that is, the incident is not influential in the decision on the order of integration.

There should be no sign of seasonal dependence in the local mean for an officially deseasonalized time series like P. It is a sign of breakdown in the official statistical approach that one does see clear evidence of seasonal mean in P.

So far in this analysis, the assumption has been that the seasonal mean is deterministic. However, the estimates of the deterministic seasonal parameters $\left(\alpha_1, \beta_1, \alpha_2 \right)$ differ somewhat in the present analysis of data for 1/85-4/05 from the values estimated in Vicente (2005) for the sample of quarterly data for 1/85-4/99 current at the time of that research. The α_1 estimate changes from -.00042, in the earlier research and sample, to -.00050, in the present research and sample, not much of a change, considering the smallest of the two estimated standard errors (.00015). The β_1 estimate changes from .00030 to

.00048, a change above the smallest of the two estimated standard errors (.00015), and the α_2 estimate changes from -.00012 to -.00021, a change also above the smallest of the two estimated standard errors (.00006). The changes are not very large, but it is well to recall that the present longer sample contains the first, except for data revisions, and only adds 24 observations to the common 60. The variation in these estimated coefficients suggests that the seasonal character of local mean might be stochastic rather than deterministic.

Model P2 specifies the seasonal mean to be deterministic at both frequencies 1 and 2. At each frequency, when the stochastic specification is applied and the model estimated, the corresponding MA parameter is estimated to be literally non-invertible. This strongly confirms the deterministic seasonal mean specification.

Consider an alternative model specification, similar to the IMA(2,1) with deterministic seasonal mean of P2, but using an AR(1) factor in place of one of the differences and introducing a parameter μ to allow for a possibly non-zero mean for the first-difference variable $\nabla \ln P$ "corrected" for the deterministic seasonal mean. Such a model allows one to test the null hypothesis of non-stationarity of this AR(1) factor and, hence, to check, from a slightly different point of view, the possibility that Model P2 is over-differenced. Call the estimated version Model P3.

$$\ln P_t = -.00050\cos(\pi t/2) + .00048\sin(\pi t/2) - .00021(-1)^t + N_t$$
$$\quad (.00015) \qquad\qquad (.00014) \qquad\qquad (.00007)$$

$$(1-.92B)[\nabla N_t - .0061] = (1-.55B)AP3_t \qquad\qquad (3.3)$$
$$(.05) \qquad\quad (.0008) \qquad (.11)$$

$$\hat{\sigma}_{AP3} = .16\%$$

Large Residuals

3/88	2.3σ
2/90	2.1σ
1/02	-2.2σ
2/04	2.4σ
3/04	-2.3σ

$r_5 = \ .16(.11)\{a_{3/03}, a_{4/04}: .04$

$r_7 = \ .16(.11)\{a_{4/88}, a_{3/90}: .06$

$r_9 = -.16(.11)\{a_{2/02}, a_{3/04}: -.06$

AP3

acf

Q (13) = 14.2

pacf

$$\overline{W}(\hat{\sigma}_W) = 0.00\ \% \ (0.02\ \%)$$

$$\hat{\sigma}_W = 0.16\ \%$$
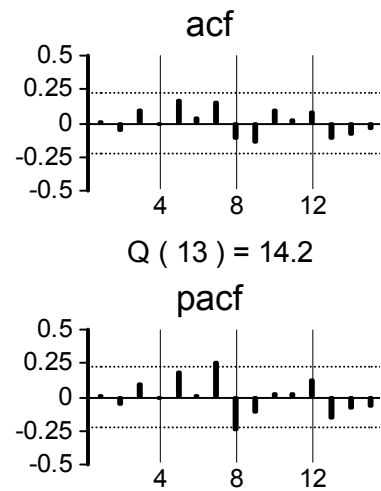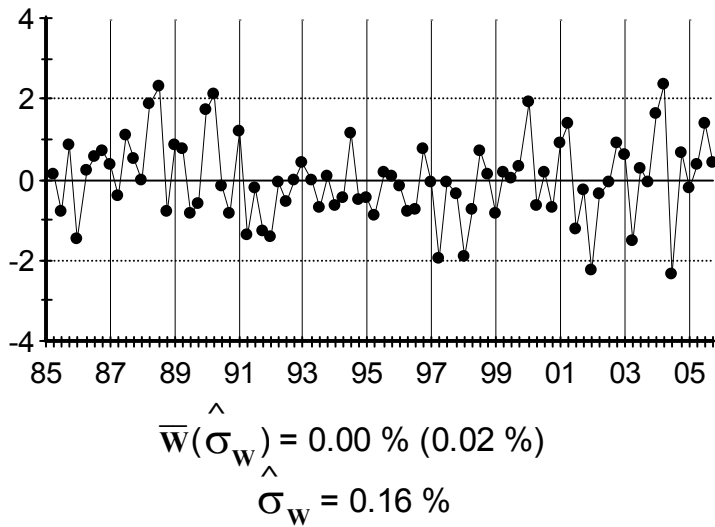
This model is very similar to Model P2 in terms of all diagnostic criteria. The SF test of $H_0 : \phi_1 = 1$ yields a test statistic of .27, much smaller than either of the critical values (1.06, 1.75) for levels of confidence (90%, 95%), so that the null hypothesis cannot be rejected. The conclusion $\ln P \sim I(2)$ is strongly confirmed.

4 Univariate analysis of the Consumer Price Index (*PC*)

The data graph of $\ln PC$ shows an upward trend with slowly varying slope. The coefficients of the *acf* die out only slowly and linearly.

lnPC

acf

Q (15) = 728.9

pacf

$$\overline{W}(\hat{\sigma}_W) = 500.09\ \% \ (1.97\ \%)$$

$$\hat{\sigma}_W = 18.02\ \%$$
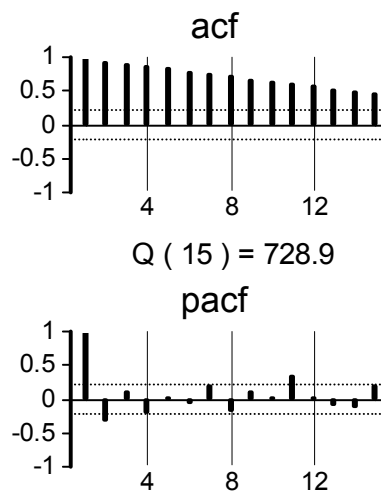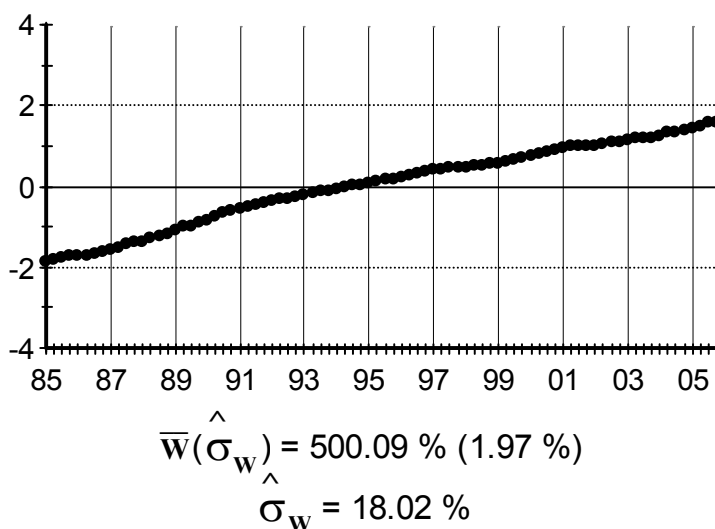
The rest of the comments made above on $\ln P$ are equally appropriate in this case. Though no definitive decision is yet taken on the order of differencing required, the evidence suggests that the level is probably stochastically non-stationary. It is thus wise to examine the series of the first difference.

$\nabla\mathbf{lnPC}$

$$\overline{W}(\hat{\sigma}_W) = 0.75 \text{ \% } (0.04 \text{ \%})$$

$$\hat{\sigma}_W = 0.39 \text{ \%}$$
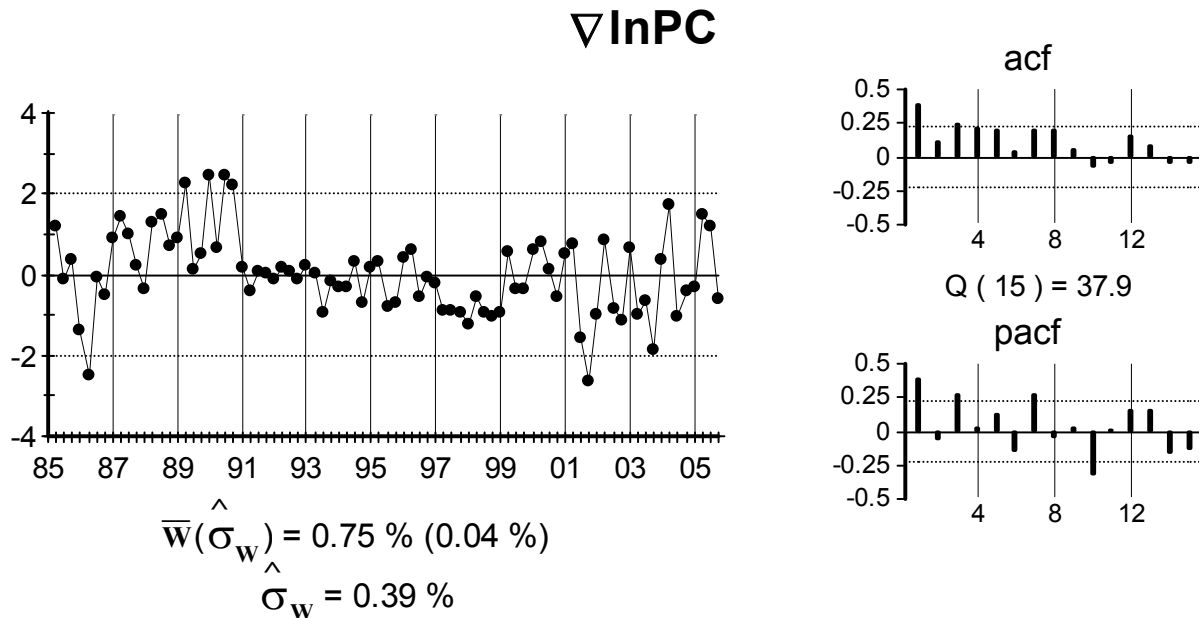
The data graph of $\nabla \ln PC$ wanders very noticeably, showing little affinity for a long-run mean value. The *acf* is dominated by positive values that die out only very slowly. This evidence suggests that even $\nabla \ln PC$ probably follows a stochastically non-stationary process. It seems wise to examine $\nabla^2 \ln PC$ before proceeding to estimate any model.

There is a noticeable reduction in variance of $\nabla \ln PC$ in the period roughly from 1/91 through 4/00. The beginning of this incident was detected in Vicente (2005), Section 3.3.3, pages 135-137, though it was not clear in that earlier analysis that this variance reduction is not permanent. An unsuccessful attempt to find an explanation for it was made; it is not apparently due to any change in methods used for constructing PC. In the present study this incident of heteroskedasticity is not treated, because any adjustment for heteroskedasticity would involve a non-linear change in the scale with which PC is measured.

The data graph of $\nabla^2 \ln PC$ appears to be well-centered, showing strong affinity for the long-run mean value of zero. The *acf* does not show signs suggestive of stochastic non-stationarity. One can conclude that $\nabla^2 \ln PC$ follows a stationary stochastic process, though it may be non-invertible. There is no sign of seasonal mean in either the data graph or the *acf*.

$$\overline{W}(\hat{\sigma}_W) = -0.01\ \%\ (0.05\ \%)$$
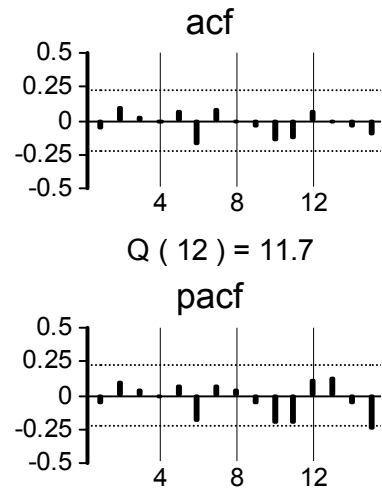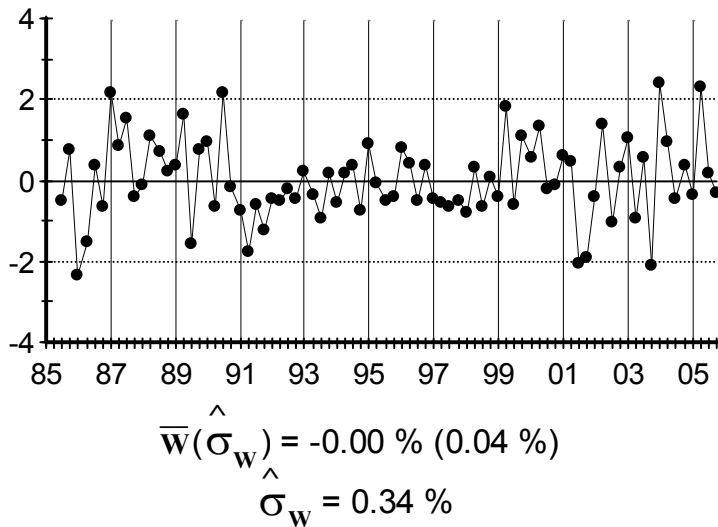
$$\hat{\sigma}_W = 0.43\ \%$$
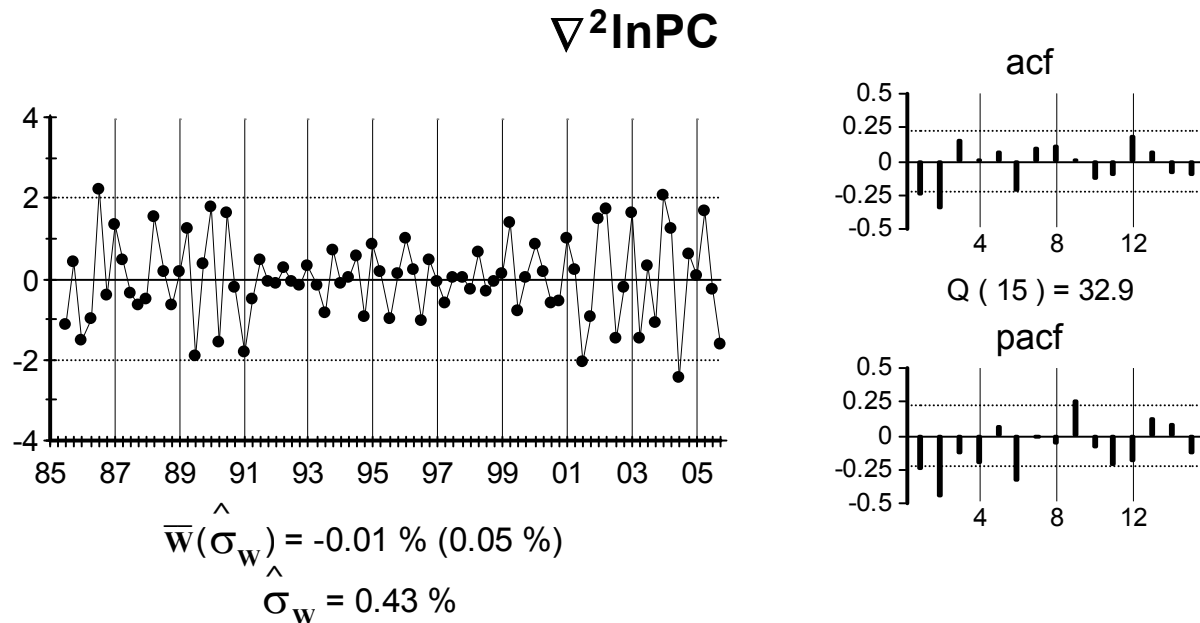
The data graph does not show either an excessively large number of extreme values or any very extreme values. Thus, intervention analysis is not appropriate at this point. The curious span of years, roughly 1991-2000, in which the variance of the series is much lower than either before or after, is evident in the data graph.

Both *acf* and *pacf* strongly suggest the presence of some MA structure, MA(2) or MA(3). The first model estimated employs the MA(3) specification and is named Model PC1.

$$\nabla^2 \ln PC_t = (1 - .67\text{B})(1 - .57\text{B})(1 + .84\text{B})\text{APC1}_t$$
$$\quad\quad (.12) \quad\quad (.12) \quad\quad (.08)$$

(4.1)

$$\hat{\sigma}_{APC1} = .34\%$$

APC1

acf

Q (12) = 11.7

pacf

$$\overline{W}(\hat{\sigma}_W) = -0.00\ \% \ (0.04\ \%)$$

$$\hat{\sigma}_W = 0.34\ \%$$

Large Residuals

1/86	-2.3σ
1/87	2.1σ
3/90	2.2σ
3/01	-2.0σ
4/03	-2.1σ
1/04	2.4σ
2/05	2.3σ

Model PC1, presented in factored form, appears to be statistically adequate. The residuals graph is well-centered, the residual mean does not differ from zero, neither *acf*, nor *pacf*, nor Q statistic indicates misspecification, and there are no very large residuals, nor is the number of large residuals very excessive. The apparent heteroskedasticity mentioned above, and not treated here, is the only noticeable flaw. The DCD statistic for the test of over-differencing is 5.50 with critical values (1.00, 1.94), indicating that over-differencing is strongly rejected. When a stationary AR(1) factor is added, the DCD statistic falls to 4.16, but this still indicates a clear rejection of the over-differencing hypothesis.
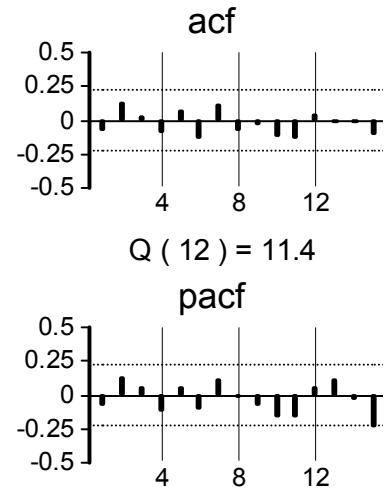
In Model PC1 there is no sign of seasonality. There is no quarter with residual values systematically of one sign, nor does the residual *acf* show relevant values in the annual lags 4, 8 and 12. Nevertheless, an over-fitting diagnostic exercise is performed and a seasonal mean

structure is discovered. The estimated Model is called PC2.

APC2
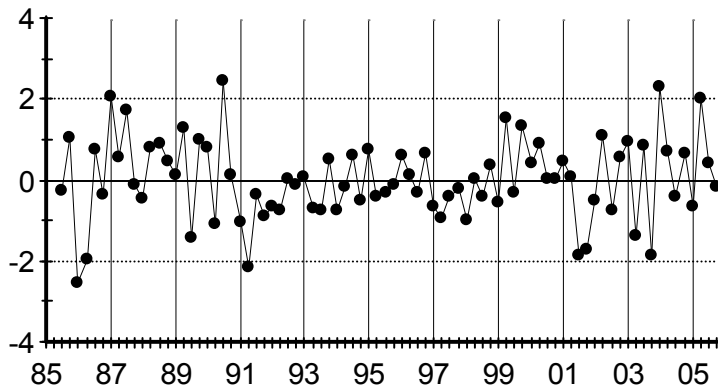
$$\overline{W}(\hat{\sigma}_W) = 0.00 \ \% \ (0.04 \ \%)$$

$$\hat{\sigma}_W = 0.33 \ \%$$

$$\ln PC_t = -.00095\cos(\pi t/2) - .00061\sin(\pi t/2) + .00001(-1)^t + N_t$$
$$\quad\quad\quad (.00045) \quad\quad\quad\quad (.00046) \quad\quad\quad\quad (.00005)$$

$$\nabla^2 N_t = (1 - .61B)(1 - .61B)(1 + .81B)APC2_t \quad\quad\quad\quad (4.2)$$
$$\quad\quad\quad (.08) \quad\quad (.08) \quad\quad (.11)$$

$$\hat{\sigma}_{APC2} = .33\%$$

Large Residuals

1/86	-2.6σ
1/87	2.1σ
3/90	2.5σ
2/91	-2.2σ
1/04	2.3σ
2/05	2.0σ

Model PC2 has all of the same diagnostic properties as Model PC1, though some deterministic seasonal mean structure is found. The DCD test for over-differencing yields the value 6.08, to be compared with critical values (1.00, 1.94), so that the over-differencing hypothesis is roundly rejected. When a stationary AR(1) factor is added to PC2, the DCD statistic falls to 5.30, but the over-differencing hypothesis is still clearly rejected. A similar result (DCD = 14.44) is obtained in a model specified exactly as PC2 but with seven

intervention parameters designed to remove possible positive distortions at lag 1.

When the parameter α_2, which is not significantly different from zero, is removed, the resulting model is practically identical to PC2, but this restriction is not employed in what follows.

A further check for over-differencing in PC2 can be performed by estimating a similar model specification, with one of the differences replaced by an AR(1) factor and introducing a parameter μ to allow for a possibly non-zero mean for $\nabla \ln PC$. The test of non-stationarity of the AR(1) factor, using the SF test, is then applied. In this case, the estimated model is practically exactly the same as PC2, but with an estimated AR(1) parameter of .88 (.08) and an estimated mean parameter of .76% (.09%). The SF statistic for the test of non-stationarity of the AR(1) factor has value .78, to be compared with the critical values (1.06, 1.75), so that non-stationarity clearly cannot be rejected. It appears very evident that $\ln PC \sim I(2)$.

Checks for stochastic seasonal mean at both frequencies 1 and 2 show clearly that the seasonal mean is deterministic in this case.

An alternative initial specification is the ARIMA(2,2,1), to be compared to PC1. The ARIMA(2,2,1) specification performs almost as well as the IMA(2,3) specification, but it suffers from several minor deficiencies: (1) the residual *acf* is not quite as clean and (2) there are clear signs that the estimation situation is not as well defined, that is, there are high correlations between the estimated AR parameters and the MA parameter. A complete sequence of analyses based on the ARIMA(2,2,1) specification confirms the presence of a deterministic seasonal mean and all tests of over-differencing confirm the conclusion that $\ln PC \sim I(2)$.

5 Univariate analysis of *P/PC*

The univariate analysis of P/PC is of interest, since both numerator and denominator are index numbers of sets of nominal prices and the ratio is a real economic variable. The identity $\nabla \ln\left(P_t/PC_t\right) \equiv \nabla \ln P_t - \nabla \ln PC_t$ reveals that the first difference of the log of the ratio is the differential between the two rates of change of numerator and denominator, themselves two alternative inflation measures. Hence the rate of change of the ratio, $\nabla \ln\left(P/PC\right)$, indicates the degree to which the two alternative inflation measures differ.

The analyses of the two previous sections establish clearly that $\nabla \ln P \sim I(1)$ and $\nabla \ln PC \sim I(1)$. In the analyses of this section the main objective is to see if $\nabla \ln\left(P/PC\right) \sim I(1)$ or if $\nabla \ln\left(P/PC\right) \sim I(0)$. In the latter case, CI(1,1) cointegration is found, and this means that it is sensible to speak of inflation as a scalar phenomenon. In the former case, the two inflation measures are not cointegrated but jointly integrated, in which case inflation must be treated as (at least) a bivariate phenomenon.

On initiating the univariate analysis of the ratio, it is known that the log of each of numerator and denominator has a deterministic seasonal mean component. Hence one anticipates the presence of a deterministic seasonal mean in the model for the ratio and it is useful to initially estimate a model with this deterministic seasonal mean included and then specify a model for the remainder.

$$\ln(P/PC)_t = \underset{(.00030)}{.00043\cos(\pi t/2)} + \underset{(.0003)}{.0011\sin(\pi t/2)} - \underset{(.00015)}{.00023(-1)^t} + \ln(P/PC)I_t$$

$$\underset{(.0003)}{\nabla\ln(P/PC)I_t + .0016} = a_t \tag{5.1}$$

$$\hat{\sigma}_a = .28\%$$

The first line of this model defines the last term, $\ln(P/PC)I$, the remainder of the level variable after removing the estimated deterministic seasonal mean terms. The second

line specifies that this remainder is taken, for present estimation purposes, to follow a random walk with drift, the latter a parameter estimated at -.16%. The parameter estimates in the first line are practically the same in an alternative estimation in which the second difference of the remainder is taken to follow an iid $N\left(0, \sigma_a^2\right)$ process. Note further that, as expected, the seasonal coefficients estimated above for the log ratio are almost exactly equal to the difference between the corresponding coefficients estimated for $\ln P$ in Model P2 and for $\ln PC$ in Model PC2. Note also that the estimated residual standard deviations in Model P2, Model PC2 and the above model respectively, are .16%, .33% and .28%; this strongly suggests that the analysis of the ratio will be dominated by characteristics of PC, because the innovation variance of this variable is about four times that of P.
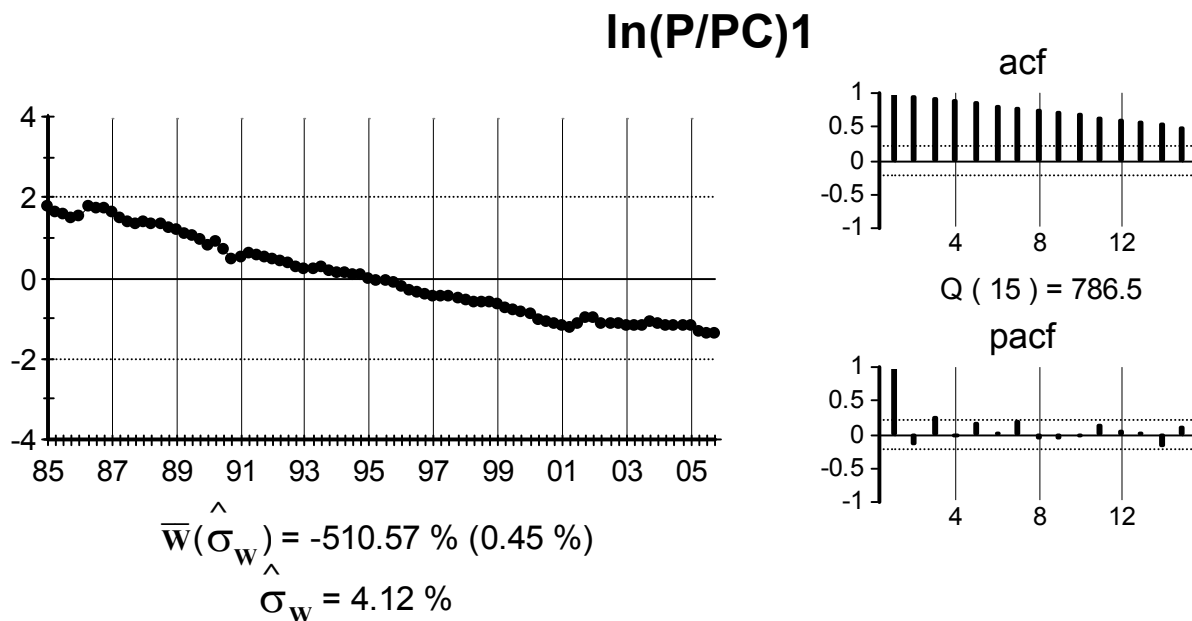
The data graph of the level of the log ratio, "corrected" for seasonal mean, shows a clear downward trend with occasional interruptions. The coefficients of the *acf* die out only very slowly and linearly. Hence it appears likely that the log ratio follows a non-stationary stochastic process. It is thus sensible to consult the first difference of the log ratio series.

$\nabla \ln(P/PC)1$

acf

Q (15) = 19.7

pacf

$$\overline{W}(\hat{\sigma}_W) = -0.16 \ \% \ (0.03 \ \%)$$
$$\hat{\sigma}_W = 0.28 \ \%$$
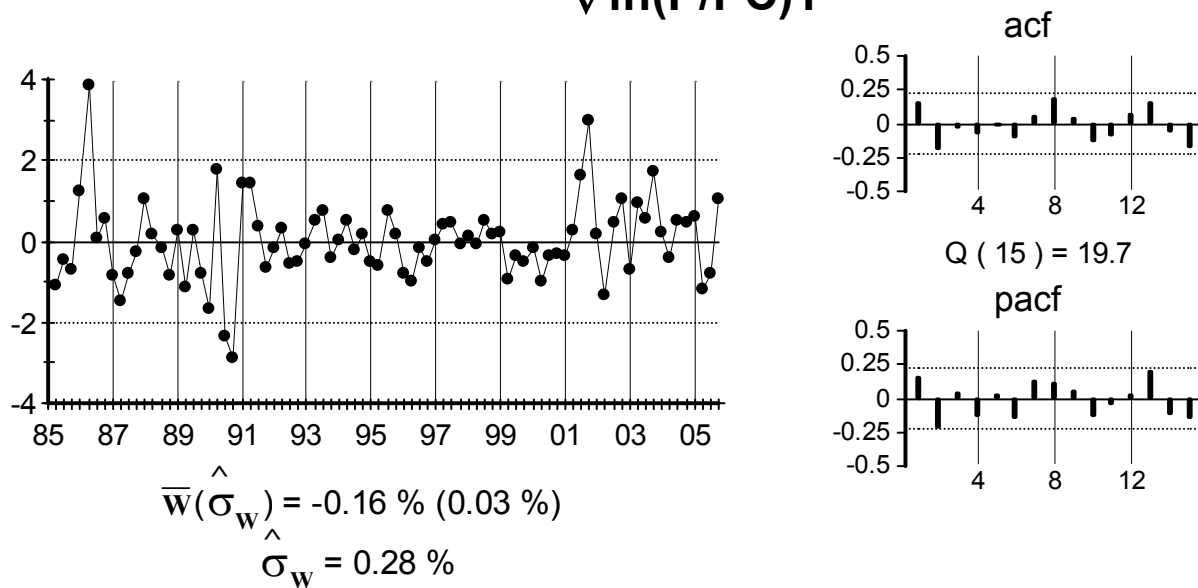
The first difference series appears to follow a stationary stochastic process: the data graph does not wander, but shows affinity for a constant long-run mean of around -.16%. The *acf* is also consistent with a stationary stochastic process, though it may be substantially distorted by extreme values: 3.9σ at 2/86, -2.4σ and -2.9σ together at 3-4/90, and 3.0σ at 4/01. Large values, isolated in time or not, tend to distort the variance measure upwards and hence to reduce all *acf* values in absolute value, obscuring whatever evidence of autocorrelation there might otherwise be. Two large values of the same sign, as occur at 3-4/90, tend to distort the first *acf* value upwards. However, no matter how distorted the *acf* may be, the data graph does suggest strongly that this series follows a stationary process. Nevertheless, it may be useful to examine the second-difference series; both $\ln P$ and $\ln PC$ have been seen to follow I(2) processes, so that the log ratio will also follow an I(2) process, unless there is cointegration, which is expected but should not be blindly assumed.

Note the very strong resemblance of the *acf/pacf* of $\nabla^2 \ln(P/PC)1$ to the case of $\nabla^2 \ln PC$ and the similarity of the data graph to the negative of that for this variable. There are, of course, some quite extreme values in this series. It is probable that the *acf* is flattened by the inflation of variance produced by these extreme values. At lag 1 of the *acf* one might

expect negative distortion, due to interactions of large values at 2-3/86 and 2-3/90, but the first of these is probably largely compensated by the positive interaction at 1-2/86, so that the net distortion may be small. At lag 2, the situation is somewhat clearer, there are strong negative interactions at 1-3/86 and 3/90-1/91, neither of which is compensated by other extreme positive interactions.

$$\overline{W}(\hat{\sigma}_W) = 0.01 \% \ (0.04 \%)$$

$$\hat{\sigma}_W = 0.36 \%$$
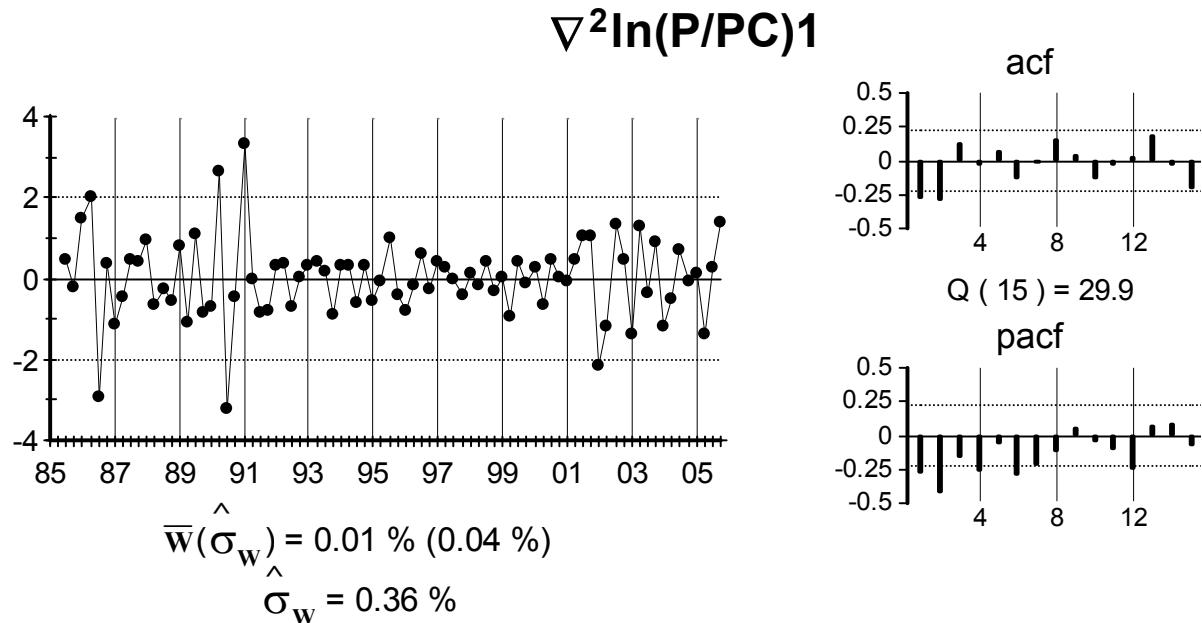
For the same reasons, based on the *acf/pacf*, that an IMA(2,3) model is initially contemplated for PC, this form is at least a useful starting point in the case of P/PC. The estimated version is Model P/PC2.

$$\ln(P/PC)_t = .00044\cos(\pi t/2) + .0011\sin(\pi t/2) - .00023(-1)^t + N_t$$
$$\quad\quad\ (.00038) \quad\quad\quad (.0004) \quad\quad\quad (.00007)$$

$$\nabla^2 N_t = (1 - .97B)(1 - .42B)(1 + .66B)AP/PC2_t \quad\quad\quad\quad (5.2)$$
$$\quad\quad (.06) \quad\quad (.14) \quad\quad (.12)$$

$$\hat{\sigma}_{AP/PC2} = .27\%$$

<u>Large Residuals</u>

2/86	3.4σ
3/90	-3.2σ
4/01	2.8σ
4/03	2.1σ

A(P/PC)2

acf

Q (12) = 8.3

pacf

$$\overline{w}(\hat{\sigma}_W) = 0.02 \text{ \% } (0.03 \text{ \%})$$

$$\hat{\sigma}_W = 0.27 \text{ \%}$$
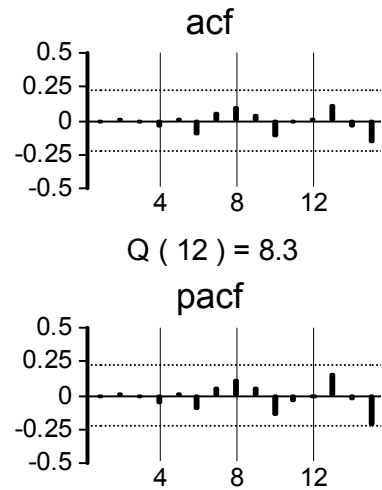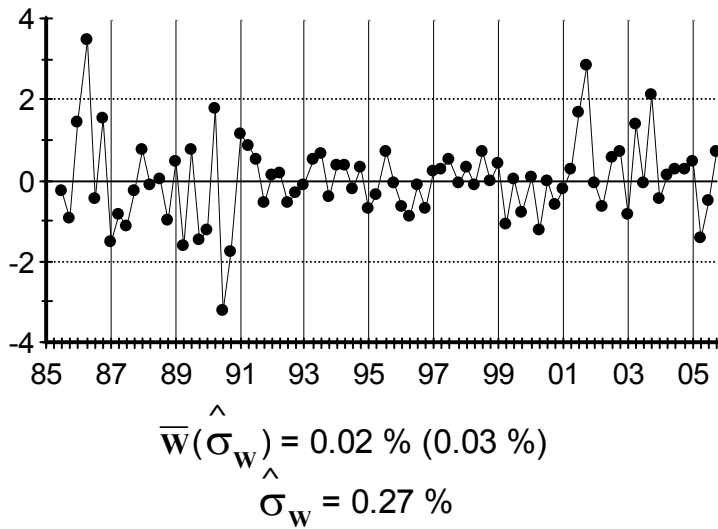
The outstanding defect of this model is that there are several very extreme residuals, suggesting quite a bit of positive distortion at lag 1, which may well influence the first MA(1) parameter upwards. The DCD statistic for testing for over-differencing is .08 with critical values (1.00, 1.94); the over-differencing hypothesis is clearly not rejected. This leads, if one ignores the real possibility that this result is due to the influences of extreme values, to a model of the form IMA(1,2) with μ, designated Model P/PC3.

A(P/PC)3
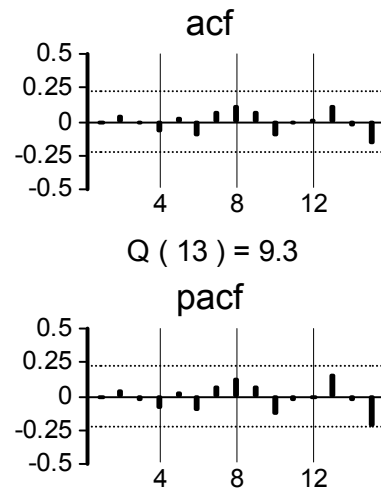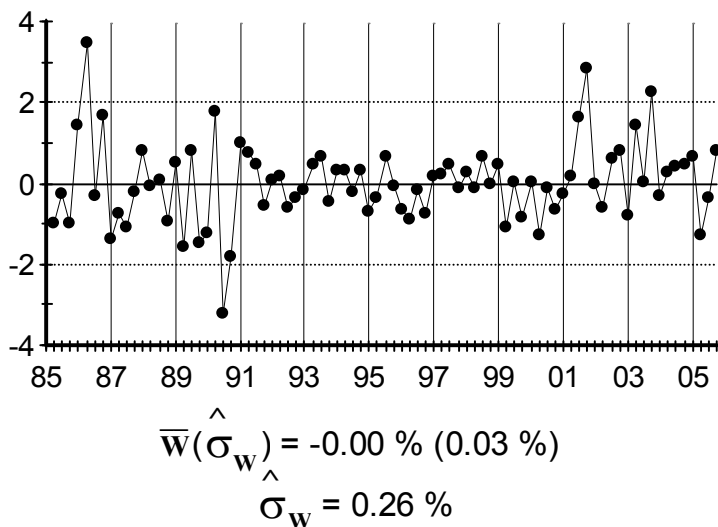
acf

Q (13) = 9.3

pacf

$$\overline{w}(\hat{\sigma}_W) = -0.00 \text{ \% } (0.03 \text{ \%})$$

$$\hat{\sigma}_W = 0.26 \text{ \%}$$

$$\ln(P/PC)_t = .00044\cos(\pi t/2) + .0011\sin(\pi t/2) - .00023(-1)^t + N_t$$
$$\quad\quad (.00038) \quad\quad\quad (.0004) \quad\quad\quad (.00007)$$

$$\nabla N_t + .0015 = (1 - .43B)(1 + .66B)AP/PC3_t \quad\quad\quad\quad (5.3)$$
$$\quad (.0003) \quad\quad (.13) \quad\quad (.12)$$

$$\hat{\sigma}_{AP/PC3} = .26\%$$

Large Residuals

2/86	3.5σ
3/90	-3.2σ
4/01	2.8σ
4/03	2.2σ

This model is very similar to the previous one and has the same main defect, namely, that there are several very extreme residuals that may well distort the *acf* sufficiently to make one imagine the model is adequate when there may be ARMA structures hidden by these distortions. In any case, if one ignores extreme values and their influences, the finding $\ln(P/PC) \sim I(1)$ is clear, that is, cointegration is found. If one adds an AR(1) factor to this model to allow for a further test of this hypothesis, the parameter turns out to be highly negative and the SF statistic rejects non-stationarity very clearly, with value 5.24 and critical values (1.06, 1.75). It is, however, essential to investigate the influences of extreme values to find out if this conclusion is or is not robust to them. For this purpose, it is useful to specify intervention terms in the absence of ARMA structure and this can be done as well with one or two differences.

To evaluate the distortions in the *acf* of $\nabla \ln(P/PC)1$ due to extreme values, a sequence of intervention models is formulated. The incident in 1-2/86 is treated with S1/86 and $s = 1$ and the incident in 3-4/01 is treated similarly with S3/01 and $s = 1$. The incident in 3-4/90 is far more complex, because, in the analysis process, further extreme values are revealed both before and after these dates, so that the incident turns out to require the introduction of S1/90 with $s = 5$, though both positive and negative effects are found at different dates in this period. These three incidents are treated separately and in all

combinations, leading finally to Model P/PC4, in which all intervention terms are present.
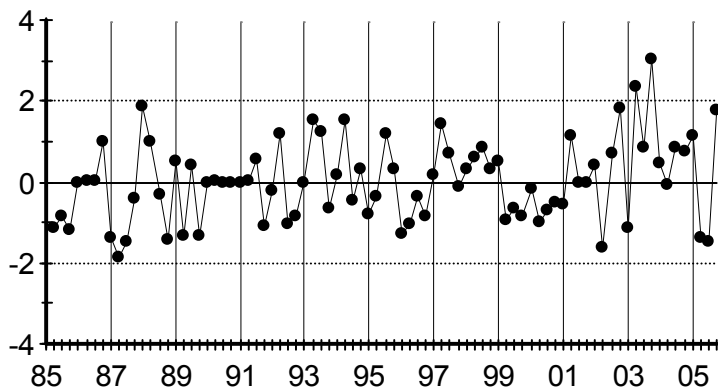
$\nabla\ln(P/PC)4$

$$\overline{W}(\hat{\sigma}_W) = -0.18\ \%\ (0.02\ \%)$$

$$\hat{\sigma}_W = 0.16\ \%$$

$$\ln(P/PC)_t = \underset{(.0011)\ (.001)}{(.0036 + .012B)\xi_t^{S1/86}}$$

$$\underset{(.002)}{\hat{g} = .016}$$
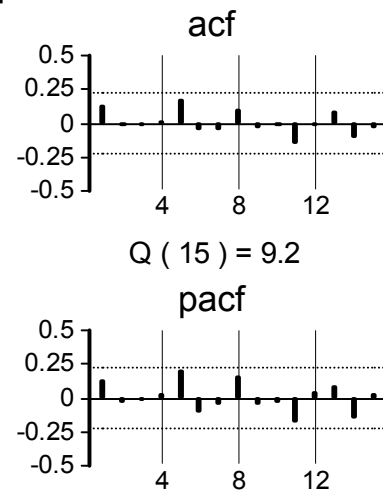
$$+\underset{(.0009)(.0012)\ (.0012)\ \ (.0004)\ \ (.0012)(.0013)}{(-.0045 + .0061B - .0067B^2 - .0081B^3 + .0042B^4 + .0051B^5)\xi_t^{S1/90}}$$

$$\underset{(.0032)}{\hat{g} = -.0039}$$

$$+\underset{(.0012)(.0005)}{(.0044 + .0084B)\xi_t^{S3/01}} +\underset{(.00018)}{.00077\cos(\pi t/2)} +\underset{(.0002)}{.0013\sin(\pi t/2)} -\underset{(.00009)}{.00036(-1)^t}\quad(5.4)$$

$$\underset{(.001)}{\hat{g} = .013}$$

$$+\ln(P/PC)4_t$$

$$\nabla\ln(P/PC)4_t +\underset{(.0002)}{.0018} = a_t$$

$$\hat{\sigma}_a = .16\%$$

Large Residuals

2/03	2.1σ
4/03	3.0σ

This model uses a large number of degrees of freedom to look for possible ARMA structure hidden by the anomalous incidents that it explicitly models. An examination of the residuals in the neighborhoods of 1-2/86, 1/90-2/91 and 3-4/01 reveals that the intervention

39

formulations appear to be adequate. Of course, the residual variance is greatly reduced in this model, making incidents, such as that visible here in 4/02-4/03, appear to be anomalous when they are not at all anomalous by the standards of the model without intervention terms. It is noteworthy that there is little evidence revealed of hidden ARMA structure in any of the diverse models estimated in this sequence with more and more intervention parameters. It is also noteworthy that the incident in 1/90-2/91 appears to be transitory, because the Student t statistic for the test that the long-run gain is zero is -.0039/.0032 = -1.22 and the critical values for 69 degrees of freedom are (1.30, 1.67). When this simplification is introduced, there are no changes in parameter estimates or in diagnostic results, though the reformed operator naturally differs.

$$
\overset{\xleftarrow{\hspace{1em}} -.80 \xrightarrow{\hspace{1em}}}{\ln(P/PC)_t} = \underset{(.0012)(.001)}{(.0037 + .012B)\xi_t^{S1/86}} + \underset{(.0014)(.0018)\ (.0018)\ \ (.002)\ \ (.0014)}{(-.0039 + .0029B - .0032B^2 - .011B^3 - .0058B^4)\xi_t^{I1/90}}
$$

$$
\underset{(.002)}{\hat{g} = .016}
$$

$$
+ \underset{(.0016)(.0009)}{(.0045 + .0084B)\xi_t^{S3/01}} + \underset{(.00019)}{.00079\cos(\pi t/2)} + \underset{(.0002)}{.0013\sin(\pi t/2)} - \underset{(.00009)}{.00036(-1)^t}
$$

$$
\underset{(.002)}{\hat{g} = .013} \tag{5.5}
$$

$$
+ \ln(P/PC)5_t
$$

$$
\nabla\ln(P/PC)5_t + \underset{(.0002)}{.0019} = a_t
$$

$$
\hat{\sigma}_a = .16\%
$$

A careful examination of the residual data graph for Model P/PC4, which is very similar to that for Model P/PC5, reveals a large number of positive interactions that would explain the small positive value in the residual *acf* without reference to distortion. This *acf* value thus reflects a small positive AR(1) parameter or a small negative MA(1) parameter. When an AR(1) parameter is added to Model P/PC5, one obtains Model P/PC6.

40

A(P/PC)6

acf

Q (14) = 8.9

pacf

$$\overline{w}(\hat{\sigma}_w) = 0.00\ \%\ (0.02\ \%)$$

$$\hat{\sigma}_w = 0.16\ \%$$

$$\ln(P/PC)_t = (.0040 + .012B)\xi_t^{S1/86}$$
$$\qquad\quad (.0016)(.001)$$

$$\hat{g} = .016$$
$$\quad (.002)$$

$$\overset{\xleftarrow{\ .72\ }\!\xrightarrow{}\ \xleftarrow{\ .75\ }\!\xrightarrow{}}{+(-.0036 + .0031B - .0029B^2 - .010B^3 - .0055B^4)\xi_t^{I1/90}}$$
$$\quad\ (.0014)(.0018)\ \ (.0020)\quad (.002)\quad (.0014)$$

$$+(.0042 + .0082B)\xi_t^{S3/01} + .00080\cos(\pi t/2) + .0013\sin(\pi t/2)$$
$$\quad (.0015)(.0016)\qquad\quad (.00018)\qquad\quad (.0002) \qquad\qquad\qquad (5.6)$$

$$\hat{g} = .012$$
$$\quad (.002)$$

$$-.00036(-1)^t + N_t$$
$$\ (.00009)$$

$$(1 - .15B)[\nabla N_t + .0019] = AP/PC6_t$$
$$\quad (.11)\qquad\quad (.0002)$$

$$\hat{\sigma}_{AP/PC6} = .16\%$$

Large Residuals

2/03	2.6σ
4/03	2.9σ
4/05	2.0σ

The AR(1) parameter is significantly different from zero and its presence allows application of the SF test for under-differencing. The SF statistic is 20.61 with critical values (1.06, 1.75), so that the hypothesis of under-differencing is clearly rejected. The conclusion

that $\ln\left(P/PC\right) \sim I(1)$ is clearly confirmed; it is not due to the influences of extreme values.

At this point, one might well wonder if $\ln\left(P/PC\right) \sim I(0)$. That is, perhaps the non-stationarity of the log ratio is entirely deterministic and the two log nominal price index variables have the CI(2,2) cointegration feature rather than CI(2,1). To check this, one can formulate a model specified exactly as Model P/PC6, but with stationary AR(2) structure, presumably with real roots, and with a linear deterministic trend component. In such a model the SF test can be applied to the largest AR(1) factor to test the null hypothesis of I(1) versus the alternative of I(0). When this is done, it turns out that $\hat{\phi}_1 = .9540 > \bar{\phi}_1 = .9524$, where the latter value is the threshold for the SF statistic. This means that the null hypothesis cannot be rejected, that is, the non-stationarity of the log ratio is found to be stochastic in nature and the CI(2,2) property is rejected in favor of CI(2,1).

6 Concluding remarks

It does not appear to the author that a consensus exists today among economists to the effect that inflation follows an $I(1)$ process in the U.S. The empirical evidence on this order of integration, cited in the introduction, is very mixed and very flimsy. Nevertheless, the empirical analyses presented in this paper, for a fairly long and recent stretch of time, and on the two measures most frequently employed for inflation, are very clear, indicating that such a consensus should reasonably be expected in the near future.

This does not mean that inflation must always follow an $I(1)$ process, either in the U.S. or elsewhere. It is not unthinkable that an effective monetary policy regime, based on feedback control and a constant, low and positive, target rate of inflation, could generate a stationary inflation time series. It may well be implausible, for many reasons, that such a regime will arise in the near future, but the results of this paper should not be construed to say that there is an inviolable law of nature making the inflation rate stochastically

42

non-stationary.

Though the belief in the univariate nature of inflation has a long tradition in macroeconomics, and no macroeconomist would lightly see it undermined, this belief is largely unconscious, not founded so far on any substantial economic theory, as far as the present author is aware, and yet today it is a testable hypothesis. This hypothesis is, furthermore, very easy to test, as this paper demonstrates. Here it is tested and very clearly not rejected. Under what conditions might it fail? So far the only known failures are to be seen in Valbuena (2002) where they arise, for a very few cases, because the rate of change of a certain sectoral nominal-price-index series does follow a stationary processes.

The analysis of section 5 reveals not only that $\nabla \ln P - \nabla \ln PC \sim I(0)$, but that the mean of this inflation differential is negative. This means that consumer price inflation is well above GDP deflator inflation in the U.S. economy in this sample. It would be of interest to know if the same result arises for the private consumption deflator from the national accounts, and whether this result is the same for all countries and periods of time or not.

The series P is found to have a seasonal mean. This is not very important quantitatively, but it does indicate a defect in the quarterly national account system, because this series is the ratio of two series that are supposedly officially adjusted for seasonality.

It is likewise very worrisome that the series PC shows a large and unexplained reduction in variance roughly in 1991-2000. This is equivalent to two non-linear changes in the measurement scale for this important economic variable, one at the beginning and one at the end of this period.

References

Andreou E, Spanos A. Statistical adequacy and the testing of trend versus difference stationarity. Econometric Reviews 2003;22(3); 217-237.

Box GEP, Tiao GC. Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association 1975;70(349); 70-79.

Box GEP, Jenkins GM, Reinsel GC. Time series analysis: Forecasting and control, 3rd ed. Prentice-Hall: Englewood Cliffs, New Jersey; 1994; 1st ed. without Reinsel, 1970.

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

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

Crowder W, Hoffman D. The long run relationship between nominal interest rates and inflation: The Fisher equation revisited. Journal of Money, Credit and Banking 1996;28; 102-118.

Crowder W, Wohar M. Are tax effects important in the long run Fisher relationship? Evidence from the Municipal Bond market. Journal of Finance 1999;54; 307-317.

Culver SE, Papell DH. Is there a unit root in the inflation rate? Evidence from sequential break and panel data models. Journal of Applied Econometrics 1997;12(4); 435-444.

Davis RA, Chen M, Dunsmuir WTM. Inference for MA(1) processes with a root on or near the unit circle. Probability and Mathematical Statistics 1995;15; 227-242.

Davis RA, Chen M, Dunsmuir WTM. 1996. Inference for seasonal moving average models with a unit root. In: Robinson PM, Rosenblatt M (Eds), Athens conference on applied probability and time series, vol. 2: Time series analysis in memory of EJ Hannan. Springer-Verlag: Berlin; 1996. p. 160-176.

Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 1979;74(366); 427-431.

Dickey DA, Fuller WA. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 1981;49(4); 1057-1072.

Evans MD, Lewis KK. Do expected shifts in inflation affect estimates of the long run Fisher relation? Journal of Finance 1995;50; 225-253.

Gallego JL. Una familia general de procesos estocásticos estacionales. doctoral thesis. Departamento de Economía Cuantitativa, Universidad Complutense de Madrid; 1995.

Gallego JL, Treadway AB. The general seasonal ARIMA family of stochastic processes. Working Paper 96.01, Departamento de Economía, Universidad de Cantabria; 1996.

Henry, ÓT, Shields K. Is there a unit root in inflation? Journal of Macroeconomics 2004;26; 481-500.

Lee H, Wu J. Mean reversion of inflation rates: Evidence from 13 OECD countries. Journal of Macroeconomics 2001;23(3); 477-487.

Ljung GM, Box GEP. On a measure of lack of fit in time series models. Biometrika 1978;65(2); 297-303.

Mauricio JA. Evaluación y maximización de la función de verosimilitud de procesos ARMA multivariantes. doctoral thesis, Departamento de Economía Cuantitativa, Universidad Complutense de Madrid; 1992.

Mauricio JA. Exact maximum likelihood estimation of stationary vector ARMA models. Journal of the American Statistical Association 1995;90(429); 282-291.

Melard G. Algorithm AS196: A fast algorithm for exact likelihood of autoregressive moving average models. Applied Statistics 1984;33 (1); 104-111.

Nelson CR, Plosser CI. Trends and random walks in macroeconomic time series: Some evidence and implications. Journal of Monetary Economics 1982;10(2); 139-162.

Ng S, Perron P. Lag length selection and the construction of unit root tests with good size and power. Econometrica 2001;69; 1519-1554.

Perron P. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 1989;57(6); 1361-1401.

Perron P. Further evidence on breaking trend functions in macroeconomic variables. Journal of Econometrics 1997;80(2); 335-385.

Rose A. Is the real interest rate stable? Journal of Finance 1988;43; 1095-1112.

Rudebusch GD. Trends and random walks in macroeconomic time series: A re-examination. International Economic Review 1992;33(3); 661-680.

Shin DW, Fuller WA. Unit root test based on unconditional maximum likelihood estimation for the autoregressive moving average. Journal of Time Series Analysis 1998;19(5); 591-599.

Valbuena Martínez, MC. Contabilidad nacional anual española: Algunas propiedades estadísticas y su interpretación económica. doctoral thesis, Departamento de Economía Cuantitativa, Universidad Complutense de Madrid; 2002.

Vicente, E. Análisis de series temporales macroeconómicas de Estados Unidos relacionadas con la inflación. doctoral thesis, Departamento de Economía Cuantitativa, Universidad Complutense de Madrid; 2005 (http://www.ucm.es/BUCM/tesis/cee/ucm-t28297.pdf).

Zivot E, Andrews E. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business and Economic Statistics 1992;10(3); 251-270.