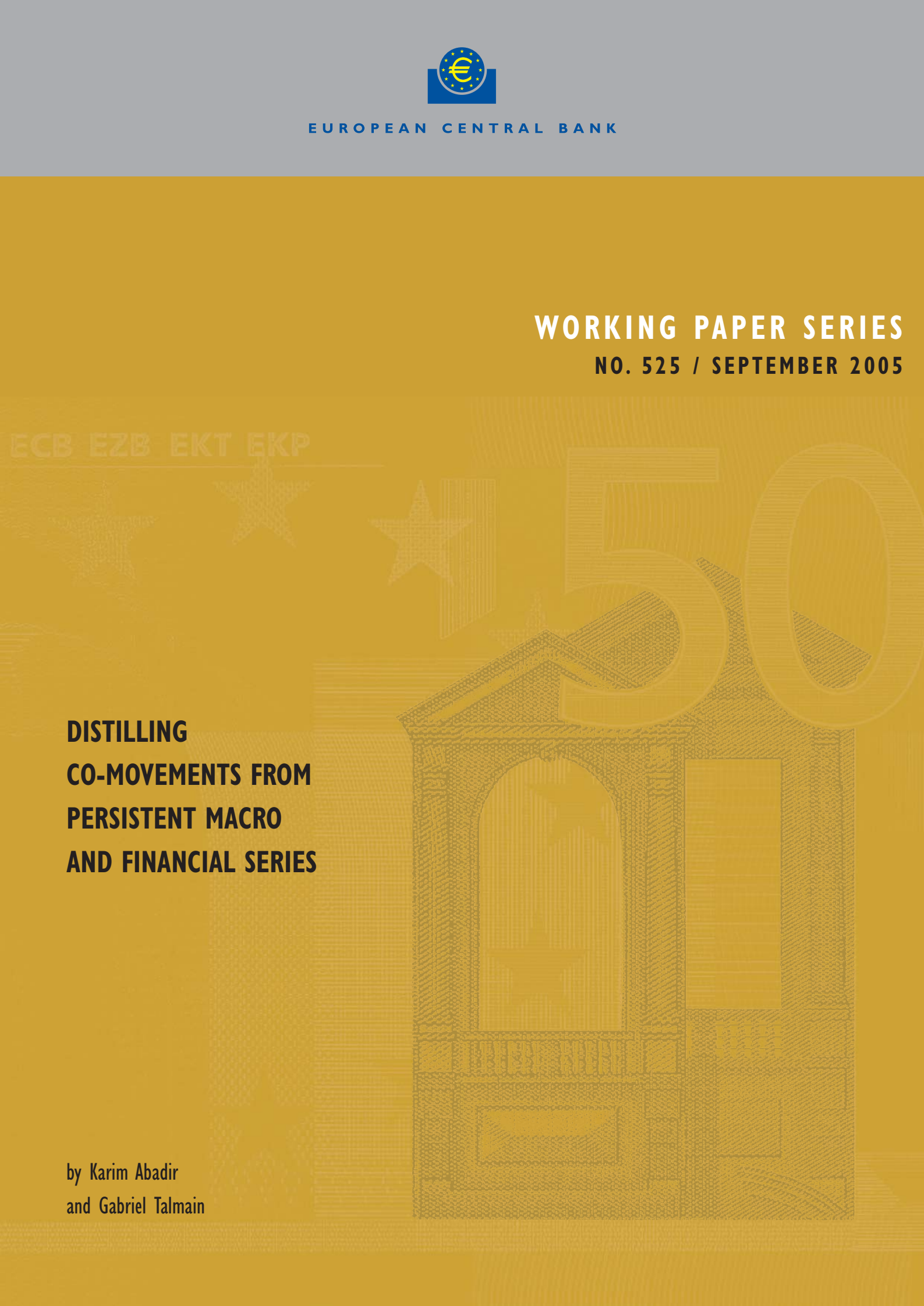

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

NO. 525 / SEPTEMBER 2005

DISTILLING CO-MOVEMENTS FROM PERSISTENT MACRO AND FINANCIAL SERIES

by Karim Abadir
and Gabriel Talmain

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES
NO. 525 / SEPTEMBER 2005

DISTILLING CO-MOVEMENTS FROM PERSISTENT MACRO AND FINANCIAL SERIES [1]

by Karim Abadir [2]
and Gabriel Talmain [3]

In 2005 all ECB publications will feature a motif taken from the €50 banknote.

This paper can be downloaded without charge from http://www.ecb.int or from the Social Science Research Network electronic library at http://ssrn.com/abstract_id=800504.

1 We thank, for their comments, seminar participants at the Bank of England, European Central Bank, and Swiss National Bank; and at Albany, Amsterdam, AUC, Bilkent, Birkbeck, Brunel, Cambridge, CORE, Edinburgh, EUI, Exeter, GREQAM, Imperial, Lausanne, Liverpool, London Met, LSE, Manchester, Maynooth, Oxford, Pompeu Fabra, Royal Holloway, Tilburg, ULB, and UCD. This paper was written, in part, while the second author was a visiting researcher at the European Central Bank. We also thank the ESRC for grant RES000230176.
2 University of York, Heslington, York YO10 5DD, United Kingdom; e-mail: kma4@york.ac.uk
3 University of York, Heslington, York YO10 5DD, United Kingdom; e-mail: gt4@york.ac.uk

© **European Central Bank, 2005**

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Internet
http://www.ecb.int

Fax
+49 69 1344 6000

Telex
411 144 ecb d

All rights reserved.

Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the author(s).

The views expressed in this paper do not necessarily reflect those of the European Central Bank.

The statement of purpose for the ECB Working Paper Series is available from the ECB website, http://www.ecb.int.

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract 4
Non-technical abstract 5
I The UIP theorem and the puzzle 8
II Uncovering the UIP puzzle: the two-step procedure 12
III The stock market application 18
IV Concluding comments 23
Appendix 23
References 27
Figures 30
European Central Bank working paper series 36

Abstract

We provide a methodology to disentangle the long-run relation between variables from their own dynamics. Macroeconomic and aggregate financial series have a high degree of inertia. If this persistence is not properly accounted for, spurious correlations will give rise to paradoxes. Our procedure shows that the Uncovered Interest Parity (UIP) puzzle evaporates when the dynamics are properly modelled: the forward premium loses all the predictive power that it seemed to have. We also show how the stock market grows in long cycles around a trend given by GDP, in a stable relation that does not break.

Key Words: ACF-based GLS procedure, Autocorrelation Function, Long memory, Nonlinearities, Uncovered Interest Parity anomaly

JEL classification: E37

Non Technical Abstract

Explaining the dynamics of national output has puzzled economists for a long time. It is crucial for the design of economic policies, that governments and central banks understand the lags with which responses to shocks and policies take effect. The timing and magnitude of their intervention depends on it.

In a previous paper, Abadir and Talmain (2002), we designed an economic model which had novel implications for the way in which these dynamics are represented. We found that national output should have the following time path: it continues on the same trajectory (either growth or stagnation) for a long time, then changes direction abruptly. This result, predicted from theory, was also validated by UK and US data. For example, the theory explains why the start (or end) of recessions and booms is so sudden, and why they are so persistent.

The premise of this paper is our recent findings, Abadir *et al.* (2005), that most macroeconomic variables, including all the variables studied in the celebrated Nelson Plosser (1982) paper, display this same type of non-linear persistence. This non-standard nature of this dynamics becomes more evident when looking at the AutoCorrelation function (ACF) of these variables. ACFs depict the decay of memory with time, as they evaluate the correlation of a series with its past. Figures 1 and 2 show the ACFs of the logarithms of the exchange rate and the S&P 500 stock market index (together with a fitting that is explained in the paper). One immediately notice the similarity in the shape of these function, and how well-structured their graphs are when compared to the time-plot of the series themselves, Figure 9 and 10. Most research in macroeconomics and in finance assumes a much simpler type of dynamic behavior based on autoregressive processes.

When conducting regression analysis on the data, there are at least two problems associated with a misidentification of the dynamic behavior. First, the regressions will purposedly uncover some relationships between variables when, in fact, no such relationship holds, i.e. one will find "puzzles". Second, even when a relationship between variables really holds, the standard

econometric analysis will produce poor results, i.e. bias, inconsistent and unstable parameter estimates.

Clearly, an appropriate tool for data analysis is needed that can handle the type of dynamic process observed in the data. In this paper, we provide such a tool and we illustrate how it will succeed where standard econometric analysis has struggled.

We provide two applications which illustrate the two problems of misidentification of dynamics. We analyze with our method the so-called 'Uncovered Interest Parity paradox' to illustrate the "puzzle" aspect. Standard regression analysis of the spot and forward exchange rates seems to imply that investors are systematically choosing the wrong currency to bet on in the foreign exchange markets. After taking this type of persistence into account, the apparent systematic error disappears!

For the second aspect, we apply our new technique to the stock market market. We find that the movements of the stock market are accounted for by the level of national output (the 'fundamentals') and by the stock market's own cyclical adjustment process (possibly coming from risk aversion). As anticipated from theory, the fundamentals prevail in determining the long-run direction of the markets, though one wouldn't have inferred that from applying standard exiting tools to the data.

The literature contains a number of paradoxes or puzzles, instances in which some very natural relationship between economic or financial variables seem to be violated. For instance, the very intuitive Uncovered Interest Parity (UIP) theorem, that should survive in a wide range of environments, has been strongly rejected. Yet, many of us feel that the theory is sound, and that what is needed is a fresh look at the statistical analysis.

It is well known that even the simplest dynamics, if not adequately accounted for, can cause biases and inconsistencies in the estimates of the relationship between the variables; e.g. see Griliches (1961) and Maddala and Rao (1973). One can see patterns in the dynamics of a series by inspecting its Auto-Correlation Function (ACF), in addition to its usual time-domain and frequency-domain depictions. Figures 1 and 2 show the ACFs of the logarithms of the exchange rate and the S&P 500 stock market index, the fitted curve in each of these graphs to be discussed momentarily. One can immediately observe how similar the shape of these two functions are, and how strong the autocorrelation is even at long lags. These well-structured ACFs are striking to anyone used to the seeing the jagged time-paths of these variables.

The strong autocorrelation means that inaccurate approximations of the dynamics are likely to lead to paradoxical results such as the UIP puzzle. Accounting for this persistence is therefore paramount. We will see that this common form of long memory is related to the one uncovered in Abadir and Talmain (2002), and cannot be approximated parsimoniously by simple linear dynamics. Integration and co-integration theory has developed into a huge new branch of econometrics to try to deal with series having a high degree of persistence, but a defining feature of integration is that it assumes that the dynamics of individual series can be approximated parsimoniously by a class of linear processes. This is why a new econometric methodology is needed to deal with this type of nonlinear long-memory.[1]

[1] A time series is said to have long memory if its autocorrelations decay very slowly, more so than the exponential rate of stationary autoregressive models but faster than the infinite memory of unit roots. Unlike the latter, long-memory series revert to their (possibly trending) means.

We exploit the common structure of these ACFs to devise a simple new method to disentangle the comovements of variables from the effects of persistence. As an illustration, we provide two applications of different natures to illustrate the versatility and potential of our method. These applications should not be misconstrued as complete modelling exercises, as such additional endeavours would be far too substantial to fit in a single study. The first application shows how our method dramatically reverses the counterintuitive finding about the forward premium in the UIP puzzle, revealing the insignificance of the premium's role. The second application is about significance, rather than the lack of it. It shows how the stock market grows in long cycles around a trend given by GDP, in a stable relation that does not break and that fits better than existing methods. Inter alia, it also predicts the rebound that followed the overreaction to the oil crisis in 1973, and the bursting of the dot-com bubble of the late 1990's.

The plan of the paper is as follows. Section I presents a baseline version of the UIP puzzle. Section II gives some details about long memory and outlines a simple two-step version of our procedure. We establish the type of persistence found in the foreign exchange rate, showing how the UIP puzzles evaporates after the dynamics of the exchange rate is fully accounted for. Section III gives an application of our fully-iterated procedure to the stock market, while the final section concludes. The full procedure and some related technical details are given in the Appendix.

I. The UIP theorem and the puzzle

One test for the efficiency of the foreign exchange market, going back to Fisher (1930), is that "speculators" will equalize the expected return on the similar short term assets across countries once converted to the same currency. However, a large number of authors analyzing the data have found systematic deviations from this norm. The data seemed to lend support to a very substantial negative relationship between the returns on holding a currency and the forward premium on it. This UIP puzzle is also known as the forward premium anomaly. Many authors have studied this very

counterintuitive finding, and excellent summaries of these are found in Lewis (1995) and Engel (1996).

This section contains two parts: the three equivalent formulations of the baseline UIP theorem, followed by the empirical puzzle. The first formulation to be given in (2), is in terms of excess returns and sets the stage for a simple graphical presentation. The second, in (3), is in terms of depreciation of a currency, and is widely used in the literature. The third, in (4), is in terms of the levels of the variables, and shows how the UIP regression can be expressed in terms of our estimation method. One will note that the three forms are derived form one another by adding or subtracting variables from both sides of the equation. Hence, the equivalence between these three forms is preserved by linear transformation of the equations. We also present the results of the first UIP regression (2), by traditional methods, to verify the presence of the puzzle in monthly data from the Bank of International Settlements (BIS).

A. *Three forms of the baseline UIP regression*

Consider a US investor who, at time t, can invest either in a domestic dollar-denominated bond or in a foreign sterling-denominated bond. The nominal interest rate, paid at maturity $t+1$, is I_t for the domestic bond and I_t^* for the foreign bond. The interest rates I_t and I_t^* are agreed upon and known at time t. Let the exchange rate be S_t, such that one pound sterling is worth S_t US dollars. Consider the following two strategies, each involving an investment of \$1 at time t:

1. Invest in the domestic bond. The yield at time $t+1$ is I_t, which can be approximated by $i_t \equiv \log(1+I_t)$ since $\log(1+I_t) \approx I_t$.

2. Invest in the foreign bond. The \$1 is first converted at the current exchange rate into £$1/S_t$. This amount is then invested in the foreign bond at time t, to produce $1+I_t^*$ at time $t+1$. Finally, converting back into dollars at the new exchange rate S_{t+1} gives $(1+I_t^*)\,S_{t+1}/S_t$, and the approximate yield is $i_t^* + s_{t+1} - s_t$, where $s_t \equiv \log S_t$ and

$i_t^* \equiv \log(1 + I_t^*)$. The difference $\Delta s_{t+1} \equiv s_{t+1} - s_t$ is the approximate rate of depreciation of the US currency.

Ignoring transaction costs, the excess return on investing in foreign asset is then defined as

$$r_{t+1} \equiv i_t^* - i_t + \Delta s_{t+1}. \tag{1}$$

The efficiency hypothesis implies that r_{t+1} should not be predictable. In particular, the forward premium $(f_t - s_t)$, where f_t is the log of the forward rate F_t, should have no explanatory power. We briefly consider three equivalent formulations of a test for this hypothesis.

The first form of the UIP regression is as follows. A direct implementation of the notion that excess returns should not be predictable is to estimate

$$r_{t+1} = \alpha + \beta(f_t - s_t) + \varepsilon_{t+1}. \tag{2}$$

Assuming for simplicity that $\mathrm{E}_t(\varepsilon_{t+1}) = 0$, the intercept α can be interpreted as the average risk premium of the domestic agent investing in the foreign bond and the slope β as the certainty-equivalent informational content of today's forward premium $(f_t - s_t)$ about the future excess return r_{t+1}. The puzzle is that the literature has found significantly negative estimates for β. A higher premium $(f_t - s_t)$ means that the forward rate indicates that the US$ is likely to depreciate, and yet the regression says that it is systematically associated instead with positive excess returns being made on the US$. The idea that agents are ready to pay more for an asset when they know it is likely to become less attractive seem to fly in the face of market efficiency or even of rationality. A positive β might have been excusable, but a very large negative one is puzzling.

The second form of the UIP regression can be derived by using the Covered Interest Parity (CIP) relation. Consider an alternative strategy for investing \$1, which is to convert it into $£1/S_t$, invest this amount in the foreign bond and sell forward the forthcoming $(1 + I_t^*)/S_t$ sterling at the forward rate F_t. Since all of these transactions can be completed today at no risk, the yield on this strategy, $i_t^* + f_t - s_t$, must be equal to the yield i_t

of investing in a domestic bond, by arbitrage. Hence, $i_t^* - i_t = -(f_t - s_t)$. By substituting this into (2), straightforward manipulations give the second equivalent form of the UIP regression:

$$\Delta s_{t+1} = \alpha + (1+\beta)(f_t - s_t) + \varepsilon_{t+1}. \tag{3}$$

Formulations (2) and (3) are equivalent, up to the CIP relation. Note that, unlike the UIP relationship, one normally observes that the CIP holds almost exactly in the major markets.

The third form of the UIP regression is obtained by recalling that $\Delta s_{t+1} \equiv s_{t+1} - s_t$ and adding s_t to both sides of (3):

$$s_{t+1} = \alpha + (1+\beta) f_t - \beta s_t + \varepsilon_{t+1}. \tag{4}$$

This formulation is in terms of the levels of the variables, with s_{t+1} as the dependent variable and only ε_{t+1} is contemporaneous to it in the equation.

B. *The UIP puzzle*

We obtained data from the BIS for the period January 1979 to February 2004. Running the regression (2) on the original data, Ordinary Least Squares (OLS) gives

$$\begin{array}{rll} \hat{r}_{t+1} = 0.00454 & -3.48\ (f_t - s_t) \\ (2.48) & (-5.25) \\ [2.42] & [-3.27] \end{array} \tag{5}$$

where t-ratios and their Heteroskedastic and Autocorrelation-Consistent (HAC) versions are given in parentheses and brackets, respectively. The RESET, a test for omitted nonlinearities, gives 5.17, with a corresponding p-value of 2% from the $F(1, 298)$ distribution. The following features stand out:

1. There is a substantial amount of autocorrelation and heteroskedasticity left over in the residuals, as is evidenced by the difference between the adjusted and unadjusted t-ratios.

2. RESET indicates substantial omitted nonlinearities in the residuals, coming from the absence of accounting for the nonlinear dynamics of the variables, as we shall see in the following sections.

3. Comments 1 and 2 point to a misspecified regression. Nevertheless, inference that is robust to *standard* autocorrelation and to heteroskedasticity can be carried out using HAC t-ratios. They show that both coefficients of the regression are significantly different from zero, the second being substantially negative, though not as much as before.

It seems that UIP is violated if one were to believe these estimated parameters. A scatter plot of the excess return r_{t+1} (on the vertical axis) versus the forward premium $(f_t - s_t)$ (on the horizontal axis) in Figure 3 tells the story. The data form a funnel shape (indicating heteroskedasticity), with a clear negative inclination ($\beta < 0$). As we shall show, these distortions are due to the graph (or regression) missing a third dimension: the nonlinear long-memory dynamics.

II. Uncovering the UIP puzzle: the two-step procedure

As mentioned in the introduction, it has been known for a long time that incomplete modelling of the dynamics gives biased and inconsistent estimates, with biases persisting even as the sample becomes infinitely large. There are two traditional solutions, and our method can be viewed as implementing a generalization of one of these, in a nonstandard way. In the following three subsections, we give a brief overview of the concepts involved, present the simple two-step version of our procedure, then apply it to the UIP puzzle.

A. *Background*

We have touched on long memory in a footnote in the introduction, and we now give a short description of it. More details are given in Beran (1992), Robinson (1994), Baillie (1996), Abadir and Taylor (1999). The ACF

$\rho_1, \rho_2, \ldots$ of a process $\{z_t\}_{t=1}^{T}$ is the sequence of correlation of the variable with its τ-th lag:

$$\rho_\tau \equiv \frac{\operatorname{cov}(z_t, z_{t-\tau})}{\sqrt{\operatorname{var}(z_t)\operatorname{var}(z_{t-\tau})}},$$

where $\rho_0 \equiv 1$. Long memory is a case where this function of τ decays very slowly as τ increases, typically hyperbolically and hence much slower than the exponential rate of decay obtained for stationary Auto-Regressive (AR) models. Unlike unit root models, shocks to a long-memory process do not have an everlasting impact.

There are two traditional treatments to deal with autocorrelated errors in the time domain. The first possibility is to use Generalized Least Squares (GLS) instead of OLS, thereby correcting the original variables for autocorrelation. For example, consider the simple illustrative model

$$\begin{aligned} z_t &= \beta x_t + u_t, \\ \text{with } u_t &= \rho u_{t-1} + \varepsilon_t \quad \text{and } \varepsilon_t \sim \text{IID}\left(0, \sigma^2\right), \end{aligned} \tag{6}$$

where $|\rho| < 1$ and ε_t are independently and identically distributed (i.i.d.) for $t = 1, 2, \ldots T$. To estimate this model, taking into account the autocorrelation of u_t, the variables of the first equation (z_t and x_t) are transformed, then they are regressed by OLS to estimate the parameter β of the relation. To illustrate, the vector $\boldsymbol{z} \equiv (z_1, \ldots, z_T)'$ is transformed into

$$\boldsymbol{A}\boldsymbol{z} \equiv \begin{pmatrix} \varphi & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & -\rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_T \end{pmatrix} = \begin{pmatrix} \varphi z_1 \\ z_2 - \rho z_1 \\ z_3 - \rho z_2 \\ \vdots \\ z_T - \rho z_{T-1} \end{pmatrix}, \tag{7}$$

where an estimate of ρ is plugged in, and where φ is usually chosen as $\sqrt{1-\rho^2}$ to stabilize the variance of the transformed residuals. The lower triangular matrix $\boldsymbol{A}$ that premultiplies the vector of z_t-values arises from the Cholesky

decomposition of the autocorrelation matrix

$$\boldsymbol{R} \equiv \begin{pmatrix} 1 & \rho & \rho^2 & \\ \rho & 1 & \rho & \ddots \\ \rho^2 & \rho & 1 & \ddots \\ & \ddots & \ddots & \ddots \end{pmatrix} = \left(1-\rho^2\right) \boldsymbol{L}\boldsymbol{L}', \tag{8}$$

where $\boldsymbol{L} \equiv \boldsymbol{A}^{-1}$. Our procedure will use Cholesky decompositions for a more general pattern of autocorrelations, as we will see.

If u_t were following an AR(p), then the lower triangular matrix $\boldsymbol{A}$ in (7) would contain $p+1$ nonzero diagonals, and the first p rows would have a normalization as was done for φ; e.g. see Chapter 5 of Amemiya (1985). When the variables have long memory, as is in our case, one needs a very large p to make this transformation. We shall see how to overcome this problem by a new method based on the ACF.

The second standard procedure is to estimate a more general model, nesting both components of (6), by using an Error-Correction Mechanism (ECM) or Autoregressive Distributed Lag (ADL) model; e.g. see Section IIIA below or Hendry et al. (1984) for an introduction. Parsimony of such a representation could be problematic if the variables have long memory, as we will show.

B. *The generalization into our two-step procedure*

We now outline the basics of our procedure. Consider the general model

$$z_t = \tilde{z}_t + u_t, \quad t = 1, 2, \ldots T, \tag{9}$$

where $\tilde{z}_t$ represents the 'fundamental' value of z_t, while u_t are the residual dynamics of adjustment towards such a value. We will assume that the variables of this equation have been de-meaned. Denoting the $T \times 1$ vector of stacked z_t values by $\boldsymbol{z} \equiv (z_1, \ldots, z_T)'$, and similarly for $\tilde{z}_t$ and u_t, we write $\boldsymbol{z} = \widetilde{\boldsymbol{z}} + \boldsymbol{u}$. We may wish to think of $\widetilde{\boldsymbol{z}}$ as the typical linear relation $\widetilde{\boldsymbol{z}} = \boldsymbol{X}\boldsymbol{\beta}$ and where $\boldsymbol{X}$ may or may not contain lagged dependent variables. Note that (6) is a special case of this model.

Now $\boldsymbol{z}$ and $\widetilde{\boldsymbol{z}}$ (hence possibly $\boldsymbol{u}$) have long memory which needs to be accounted for. A GLS-type procedure is to:

1. Cholesky-decompose the autocorrelation matrix of $\boldsymbol{z}$, namely

$$\boldsymbol{R} \equiv \begin{pmatrix} 1 & \rho_1 & \rho_2 & \ddots & \rho_{T-2} & \rho_{T-1} \\ \rho_1 & 1 & \rho_1 & \ddots & \ddots & \rho_{T-2} \\ \rho_2 & \rho_1 & \ddots & \ddots & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots & \rho_1 & \rho_2 \\ \rho_{T-2} & \ddots & \ddots & \rho_1 & 1 & \rho_1 \\ \rho_{T-1} & \rho_{T-2} & \ddots & \rho_2 & \rho_1 & 1 \end{pmatrix},$$

   into $\boldsymbol{R} = \theta \boldsymbol{L}\boldsymbol{L}'$, where $\boldsymbol{L}$ is lower-triangular. The scalar (1×1) factor of proportionality θ is defined by making the last element of $\boldsymbol{L}^{-1}$ equal to 1, just as with $\boldsymbol{A} \equiv \boldsymbol{L}^{-1}$ of (7). In the example of (8), we get $\theta = 1 - \rho^2$. See the Appendix for further details.

2. Then, the transformed data $\boldsymbol{A}\boldsymbol{z}$ and $\boldsymbol{A}\widetilde{\boldsymbol{z}}$ can be regressed by traditional methods. When $\widetilde{\boldsymbol{z}} = \boldsymbol{X}\boldsymbol{\beta}$, the transformation $\boldsymbol{A}\widetilde{\boldsymbol{z}}$ boils down to transforming $\boldsymbol{X}$ into $\boldsymbol{A}\boldsymbol{X}$.

The Cholesky decomposition command is built-in as standard in all matrix-handling languages, such as Gauss and Matlab.

In the case of an AR(1), we have $\rho_\tau = \rho^\tau$, so knowledge of ρ allows us to fill the whole $\boldsymbol{R}$ matrix. However, in the general case, estimating $\boldsymbol{R}$ requires estimating $T-1$ parameters ρ_τ, which is tantamount to infeasible GLS. Our solution is to fit the ACF of $\boldsymbol{z}$, using a variant of the functional form in Abadir and Talmain (2002)

$$\rho_\tau \approx \frac{1 - a\left[1 - \cos\left(\omega\tau\right)\right]}{1 + b\tau^c}, \tag{10}$$

with only 4 parameters to estimate rather than $T-1$. The analytical justification for this form is now given, with the empirical evidence following immediately afterwards.

Abadir and Talmain (2002) showed that GDP and other variables driven by it in their model follow a new type of mean-reverting long-memory process, though their econometric context was that of a reduced form for a single variable in the ACF domain, rather than the new idea of incorporating ACFs into multivariate time-domain estimation. The shape of the ACF of this long-memory process, derived from an economic model, was very different from those predicted by traditional Auto-Regressive Integrated Moving-Average (ARIMA) models, including the linear fractional I(d) case. The empirical ACF of GDP for the US, UK and other developed nations turned out to have the shape predicted by our theoretical derivations. The functional form of the ACF in Abadir and Talmain (2002) is just the denominator of (10), which controls the decay of memory, and we have added here the numerator. The justification for this is that we need to consider rates, as well as levels of variables, and the ACF of Abadir and Talmain (2002) is only valid for the latter. The numerator accommodates short-term cycles. When $a = 0$ or $\omega = 0$, we are back to the old form of the ACF, so the modification is compatible with variables in either type of format: levels or rates.

C. *Application of our ACF-based procedure to the UIP puzzle*

Using nonlinear LS, we fit (10) to the US$-UK£ exchange rate data that we have. We get the estimate

$$\hat{\rho}_\tau = \frac{1 - 0.99\left[1 - \cos\left(0.0517\tau\right)\right]}{1 + 0.0118\tau^{0.982}}. \tag{11}$$

The original and fitted data are depicted in Figure 1, and reveal a striking accuracy of this simple functional form for the ACF. The shape of the empirical ACF reveals the presence of substantial memory. Perceptive traders may rationally take into account the persistent nature of the former series. Consequently, one must take into account the persistence in the time series before attempting to run a regression such as (2).

After transformation of the data in (2), including the constant, we obtain

the following estimates

$$\begin{array}{lll} \hat{r}_{t+1}^{\text{acf}} = 0.0000250 & -1.35 \left(f_t^{\text{acf}} - s_t^{\text{acf}}\right), \\ (0.01) & (-0.60) \\ [0.03] & [-0.48] \end{array} \tag{12}$$

where the superscript "acf" denotes the transformed variables. The RESET for omitted nonlinearities is now 0.43, with a p-value of 51% indicating that there is no leftover nonlinearities. We can see that the HAC adjustment makes almost no difference to the t-ratios now, so there is very little residual dynamics or heteroskedasticity leftover. Both estimated coefficients are insignificant, as theory had predicted. The transformation has changed the story in a dramatic way. Once the dynamics of the problem are properly taken care of, the forward premium $(f_t - s_t)$ has no predictive power at all for the excess returns r_{t+1}, neither linearly (because $\hat{\beta}$ is insignificant) nor as a nonlinear function (because the RESET statistic is insignificant).

The contrast is even more striking if we compare the bivariate scatter plots of r_{t+1} and $(f_t - s_t)$, before and after the ACF-based transformation, in Figures 3 and 4, respectively. With the original data, there seemed to be a clear negative relation between r_{t+1} and $(f_t - s_t)$. After the transformation, it looks like a round 'cloud' with no particular tendency or clear deformation, which is the way that the scatter plot of two independent variables should be, after standardization of the scale for both axes. The relation found at the beginning between these variables was spurious, and was just an artifact of the long memory of the variables involved in the regression, and the incomplete modelling of these dynamics. Of course, we could improve our results further by including risk premia, transaction costs and/or peso problems. What we have done here is to show that our simple dynamics already provide a lot of explanatory power for the series. Implementing the fully iterative ML procedure of the Appendix, we get comparable results to the two-step GLS of this section. For example, instead of (11), we get

$$\hat{\rho}_\tau = \frac{1 - 1.07\left[1 - \cos\left(0.0541\tau\right)\right]}{1 + 0.0218\tau^{0.949}}$$

for the ACF.

Engel and Hamilton (1990), Diebold et al. (1994), Baillie and Bollerslev (1994, 2000) have presented evidence of long swings and long memory in exchange rates, but have not used (11). Once the nonlinear long memory is accounted for, the forward premium loses all of its predictive power. This nonlinear long memory is close to a I(1) process, as shown in Abadir and Talmain (2002, p.765). Indeed, if the set of allowable processes is restricted to autoregressive models, then the near random walk process of Engel and West (2004) becomes the most accurate approximation.

III. The stock market application

In a monopolistically competitive economy, firms will make an economic profit, which can be thought of as dividends in a Modigliani-Miller setting. In equilibrium, the value of a firm is equal to the (stochastically) discounted flow of its dividends. However, these dividends are not completely exogenous in a general equilibrium framework. Talmain (2003) showed that, on a balanced growth path, the capitalization of stock market should be proportional to GDP.

In the second application of our estimation procedure, we show that the S&P 500 index does not have the unit root often debated in the literature on stock market efficiency. Rather, it grows in long cycles guided by the trend line given by GDP, as seen in Figure 5 where both variables are in logarithms of real US$. There are precursors to both econometric components of our result. De Bondt and Thaler (1985, 1987) first found long cycles in individual stocks, and more recently Cooper et al. (2004) give a long subsequent reference list; while Cavaliere (2001) devised a nonparametric test (but no model) that detected the presence of long memory in S&P 500 and rejected the unit root hypothesis.

In this section, we study this relation by two different methods: the standard one (unit roots and co-integration analysis) and our procedure. We also show that standard methods trying to model S&P 500 would end up with estimated parameters that are unstable, often oscillating as more data becomes available, trying to mimic the nonlinearity of the process. S&P is

therefore not an integrated process, and so it is not valid to use co-integration as a tool for analyzing this subject.

A. *The estimates from standard analysis*

Define y_t and s_t to be the logarithms of the annual real GDP and of the real stock market index, respectively.[2] Consider the error-correction model of s_t in terms of y_t,

$$\begin{aligned}(13)\ \Delta s_t \ &= \ \alpha + \left(\beta_1 \Delta s_{t-1} + \cdots + \beta_p \Delta s_{t-p}\right) + \left(\gamma_1 \Delta y_{t-1} + \cdots + \gamma_q \Delta y_{t-q}\right) \\ &\quad -\delta\left(s_{t-1} - y_{t-1}\right) + \delta_1 y_{t-1} + \varepsilon_t,\end{aligned}$$

where the are no contemporaneous variables on the right-hand side, apart from ε_t. If $\delta \neq 0$, the error-correction mechanism is defined by the expression

$$-\delta\left(s_{t-1} - y_{t-1}\right) + \delta_1 y_{t-1} = -\delta\left(s_{t-1} - \lambda y_{t-1}\right),$$

where $\lambda \equiv 1 + \delta_1/\delta$. The ECM represents the long-run relationship between s and y in 'equilibrium': $s_{\mathrm{e}} = \lambda y_{\mathrm{e}}$. The long-run proportionality between s_{e} and y_{e} can be investigated by testing the hypothesis H_0: $\delta_1 = 0$. Let d_{t-1} be the deviation at time $t-1$ of s_{t-1} from its long-term value s_{e}, that is, $d_{t-1} \equiv s_{t-1} - s_{\mathrm{e}}$. This deviation will pull s_t back towards its long-term equilibrium value by δd_{t-1} when $\delta > 0$. A small δ indicates a weak tendency for return to the long-term equilibrium.

Hendry and Von Ungern-Sternberg (1980) generalize this model to include an Integral Correction Mechanism (ICM), where cumulative imbalances of

[2] We use the S&P 500 as our stock index. The other variables we need are an aggregate price index and real output which, unlike a stock index, is only available at low frequency. To avoid this seasonality problem interfering with the results, we use annual data. Real annual GDP is available from National Accounts. The consumer price index CPI is also available from the same source at a monthly frequency. Theoretically, the value of firms is related at each moment in time with current output. However the figure for annual GDP is the production throughout the year. Hence, it must be related to an average stock index over the year. We first construct the average daily close of the S&P 500 at the highest frequency at which the CPI is available. Next, we divide this average by the CPI to obtain an index of the real value of the stock market. Finally, we convert this real stock market index into a yearly index.

$(s_{t-1} - y_{t-1})$ play a role in the catching-up of s_t with its 'equilibrium' value. This is an important mechanism. When equity prices increase faster than their fundamental value, they build up the wealth of the stockholders faster than on a balanced growth path. This wealth buildup creates the condition for a tendency-reversal when stockholders try to convert their capital gains into additional discretionary spendings. Other factors, such as capital accumulation, pull in the same direction. We summarize all of these factors in one variable, the adjustment overhang, which is the cumulative sum of departures of the log of S&P from its fundamentals. Define

$$x_{t-1} \equiv \sum_{j=1}^{t-1} (s_j - y_j) \,.$$

Adding this explanatory variable to the right-hand side of (13), we also need to add a balancing linear trend, in case the sample did not start with $(s_0 - y_0)$ being at its equilibrium value of zero.

For S&P 500 over 1953-2003, we obtained the regression

$$\begin{array}{rcllll} \widehat{\Delta s_t} & = & 0.00514\ \Delta s_{t-1} & -0.387\ \Delta s_{t-2} & -0.331\ \Delta s_{t-3} & \\ & & (0.04) & (-2.67) & (-2.21) & \\ & & -0.0456\ x_{t-1} & -0.00289\ t & +1.617\ \Delta y_{t-2}, & \\ & & (-4.73) & (-2.30) & (1.81) & \end{array}$$

where the t-ratios are in parentheses. Insignificant variables have been dropped (except Δs_{t-1} which we will discuss in the next paragraph), and we do not report the constant.

The regression indicates that the proportionality hypothesis holds, and that the ICM plays a more important role than the ECM. The fit is $R^2 = 37.4\%$. Closer inspection reveals the fragility of the estimates. In Figure 6, the central lines present recursive parameter estimates as the sample is increased to its full size, with ± 2 standard-error bands for approximate 95% confidence intervals. Panels 1-6 represent the parameters of the variables in the order that they appear in the regression. For stability of the parameter estimates, the central lines should be nearly horizontal. Significance of these estimates occurs when the bands do not include zero. The bands for the

initial estimates are understandably large, since these are based on the very few first observations. Otherwise, we make the following two observations.

First, although most parameter estimates are significant, this is not consistently so throughout the period. A sample stopping a few years short would have found the first parameter significant, with the third and fifth insignificant. The sixth is marginal throughout, depending on the significance level that is chosen. Second, the parameter estimates are oscillating, trying to replicate cycles in the nonlinear dynamics of the series.

Of course, one can include more lags of the variables in question, but this worsens rather than improves stability, while not improving the fit since these additional estimates are insignificant. Also, the RESET is 2.69, with a borderline p-value of 10.8%, indicating potentially some omitted nonlinearities.

B. *The estimates from our ACF-based ML procedure*

Once we look at the ACF of s_t, it becomes clear that s_t is not a unit-root process, but rather the nonlinear long-memory type discussed earlier. Fitting to it our theoretical functional form (10), we get the fitted curve in Figure 2 where the approximation is again very good. However, because the right-hand side of the relation is not zero in this application, the best fit will be obtained from the fully iterative (rather than the two-step) procedure.

We now use the fully iterative ML procedure on the same period. We get the joint estimates for the adjustment dynamics of u_t as

$$\hat{\rho}_\tau = \frac{1 - 1.041\left[1 - \cos\left(0.612\tau\right)\right]}{1 + 0.169\tau^{1.109}},$$

and for the corresponding transformed variables

$$\begin{aligned}\widehat{\Delta s_t^{\mathrm{acf}}} \;=\; & \underset{(-5.59)}{-0.708\ \Delta s_{t-1}^{\mathrm{acf}}} \quad \underset{(-4.21)}{-0.586\ \Delta s_{t-2}^{\mathrm{acf}}} \quad \underset{(-4.66)}{-0.530\ \Delta s_{t-3}^{\mathrm{acf}}} \\ & \underset{(-5.63)}{-0.0690\ x_{t-1}^{\mathrm{acf}}} \quad \underset{(-2.80)}{-0.00452\ t^{\mathrm{acf}}} \quad \underset{(-2.48)}{-1.49\ \Delta y_{t-1}^{\mathrm{acf}}}.\end{aligned}$$

Insignificant variables have been dropped, and the proportionality hypothesis is accepted as before. In contrast to the unit roots and co-integration analysis,

the coefficients are now very stable, throughout the sample, as can be seen in Figure 7. The fit in terms of the original data Δs_t is $R^2 = 50.8\%$, and the p-value for the improvement in fit compared to the standard model is 3.91%, a substantial improvement. Also, the RESET is now 0.10, with a p-value of 75.6%, indicating clearly no leftover nonlinearities. The previously omitted nonlinear dynamics of S&P 500, away from the long-run proportionality to GDP, have been accounted for by the ACF dynamics. In the context of Figure 5, GDP provides the stochastic trend line around which S&P 500 moves in long-memory cycles. Notice that these cycles have a pattern that is fairly well behaved, in contrast with the unpredictability that a unit-root model would have implied. As time passes, the variance bounds of a random walk would diverge linearly away from its trend line, and there would be no tendency for mean reversion (zero probability of this happening). However, in Figure 5, we see no such feature, neither for GDP nor for S&P. The analysis of this section confirms it for S&P. For GDP, see Abadir and Talmain (2002).

Figure 8 shows the implied predictions of the percentage changes in S&P 500. The actual changes are generally well reflected in our fitted values. There are a few notable exceptions. The excess volatility of S&P in the early 70's is detected by our predictions pointing to less pronounced changes. The result was that the market undertook a correction, immediately the next period, to come nearer the fundamental values from our equation. A similar story can be told about the end of the 1990's and early 2000's. Our prediction was pointing the way down at the heydays of the dot-com bubble and, as it burst so dramatically, our method showed that the fall of the market has been overdone. A subsequent upward correction took place, as anticipated by our prediction.

Finally, notice that in both applications of our method, once the long cycles are taken into account, the residuals were found to contain no structural breaks. We did not need to add to the models any dummy variables for breaks and/or extreme events, even in periods where the relation in terms of the raw data seemed at first sight to break, e.g. due to Britain's exit of the exchange-rate mechanism or to extended S&P bubbles or corrections. The ICM plays the following two roles. First, it makes sure that if S&P strays

from GDP too fast, it is brought back faster than usual. This can be seen in the last four years, which correspond to the bubble that burst, and where the cycle was shorter than the previous ones. Second, and as a result, the ICM ensures that the S&P cycles are of random length (or frequency), once the two dynamics arising from ICM and ACF are added up.

IV. Concluding comments

Integration and co-integration have had a huge impact on the analysis of macroeconomic and aggregate financial data. It was a good first step in establishing methods to deal with potentially nonstationary variables. Here, we present an alternative econometric method of analysis that is justified by the macroeconomic model of Abadir and Talmain (2002), and we show how it explains the evolution of exchange rates and S&P 500. More generally, our method has the potential to reveal new insights when two conditions are satisfied: whenever a model requires us to disentangle the dynamics of a dependent variable from its relationship with other variables of interest, and when these dynamics are of the long-memory form. The first condition is the norm. The second feature is increasingly encountered, given the recent econometric advances in handling long memory processes and the evidence it has so far uncovered. For example, see the arguments in Abadir and Taylor (1999) and the numerical results in Diebold and Rudebusch (1989), Baillie and Bollerslev (1994, 2000), Gil-Alaña and Robinson (1997), Chambers (1998), Abadir and Talmain (2002).

APPENDIX

Here we give details on the GLS procedure that was outlined earlier in its basic two-step form. To complement GLS, we also present a quasi Maximum Likelihood (ML) procedure. The adjective "quasi" is standard, and indicates that the likelihood function is based on the assumption of asymptotic normality of the distributions.

To simplify the exposition, we adopt the linear model

$$\boldsymbol{z} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \text{ with } \boldsymbol{u} \sim \mathrm{D}_1\left(\boldsymbol{0}, \boldsymbol{\Sigma}\right) \tag{14}$$

where $\boldsymbol{\Sigma}$ is the $T \times T$ autocovariance matrix. We assume that u_t $(t = 1, 2, \ldots, T)$ is mean-reverting, so that $\boldsymbol{\Sigma}$ has a symmetric Toeplitz structure. Then, $\boldsymbol{\Sigma} \propto \boldsymbol{A}^{-1}\boldsymbol{A}'^{-1}$, where $\boldsymbol{A}$ is the lower triangular matrix that removes autocorrelation from the data, taking the form

$$\boldsymbol{A} \equiv \boldsymbol{L}^{-1} = \begin{pmatrix} * & \mathbf{0} \\ -\boldsymbol{\alpha}' & 1 \end{pmatrix}, \text{ with } \boldsymbol{\alpha}' \equiv (\alpha_{T-1}, \cdots, \alpha_2, \alpha_1),$$

and where $\boldsymbol{L}$ is the matrix seen in Section IIB and obtained from the normalized Cholesky decomposition of the autocorrelation matrix $\boldsymbol{R}$. Therefore,

$$\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \text{ with } \boldsymbol{e} \sim \mathrm{D}_2\left(\mathbf{0},\ \sigma^2\boldsymbol{I}_T\right)$$

where σ^2 does not depend on the other parameters of the model and, if D_1 is the normal distribution, then D_2 is normal too. The vector $\boldsymbol{e}$ of transformed residuals is now i.i.d., and standard estimation procedures can be applied to the transformed model. Iteration between the two steps gives GLS. Alternatively, the criterion $(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})$ can be minimized for all parameters jointly.

The ML estimators are obtained by maximizing

$$-\log\left(\det\left(\boldsymbol{\Sigma}\right)\right) - (\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta}),$$

where only the first term differs from the GLS criterion, and it has the beneficial effect of ensuring that the elements of the diagonal of $\boldsymbol{A}$ are not too far from unity, which makes e_t like ε_t of (6). Concentrating the log-likelihood with respect to

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{z} \text{ and } \hat{\sigma}^2 = \frac{1}{T}\left(\boldsymbol{z} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)' \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{z} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)$$

gives

$$(15) \qquad -\log\left(\det\left(\left(\boldsymbol{z} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)' \boldsymbol{R}^{-1} \left(\boldsymbol{z} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right) \boldsymbol{R}\right)\right) - T\left(1 - \log\left(T\right)\right)$$

to be optimized with respect to the parameters of the ACF: the optimization now depends only on the four parameters in (10) that determine the autocorrelation matrix $\boldsymbol{R}$. The term $T\left(1 - \log\left(T\right)\right)$ is just a constant that does not affect the optimization, and can be dropped from the criterion.

We note the following requirements and/or features of $\boldsymbol{R}$ and the corresponding $\boldsymbol{A}$:

1. In estimating the parameters of the ACF, one needs to restrict their values so that $\boldsymbol{R}$ is positive definite. There is no explicit formula for this restriction (because no explicit solution exists for the roots of polynomials of order greater than 4), but it is straightforward to implemented numerically.

2. Fitting the early values of ACF (τ near zero) is more important than the rest, which is based on fewer observations. For this reason, we only use the first half of the ACF plot to estimate its 4 parameters. If necessary, this may be fine-tuned by a weighting function less abrupt than the one we effectively use, but the optimal estimation of such weights is beyond our scope here.

3. The lower triangularity of $\boldsymbol{A}$ ensures that each element of the transformed $\boldsymbol{z}$ is constructed only from past and current values of z_t; e.g. see (7). The same comment applies to $\boldsymbol{X}$.

4. The elements in the last row of $\boldsymbol{A}$ have an interpretation as the coefficients of an AR($T-1$) representation for the last transformed data point. One may wish to restrict the optimization procedure such that it produces stationary roots for this AR representation, but we have not done so. Note that *any* non-explosive process, whether nonlinear and/or nonstationary, can be represented as an invertible MA having time-varying coefficients, which explains the time-varying AR representations implied by the rows of $\boldsymbol{A}$. This is Cramér's decomposition of time series, a generalization of Wold's decomposition, and explains how the nonlinear process of Abadir and Talmain (2002) can be estimated by our linear representation procedure.

5. As with all GLS procedures, the initial values have to be treated with care. For example, in (7), one has a separate treatment for the trans-

formed z_1. The problem is exacerbated with long memory. One simple rule of thumb is to discard the first couple of transformed data points.

6. A well-known feature of GLS is that the constant, once transformed by any $\boldsymbol{L}^{-1}$, is not a constant vector anymore; e.g. use $\boldsymbol{\iota} \equiv (1, \ldots, 1)'$ instead of $\boldsymbol{z}$ in (7) and compare the first element to the remaining $T-1$. In our GLS procedure, it is therefore assumed that the data ($\boldsymbol{z}$ and $\boldsymbol{X}$) have been de-meaned before being transformed. This is because the procedure is based on transforming vectors, say $\boldsymbol{y}$, centered around $\mathbf{0}$ from $\boldsymbol{y} \sim \mathrm{D}_1(\mathbf{0}, \boldsymbol{L}\boldsymbol{L}')$ into $\boldsymbol{L}^{-1}\boldsymbol{y} \sim \mathrm{D}_2(\mathbf{0}, \boldsymbol{I}_T)$. Having a nonzero mean in $\boldsymbol{y}$ would have introduced a common factor of $\boldsymbol{L}^{-1}\boldsymbol{\iota}$ in all these transformed variables, which may dominate these series and produce some seemingly common factor. If a constant is required in the regression, it should be transformed separately then added to the regression for transformed variables, as we have done in the UIP example.

In Section II we chose GLS, in preference to OLS for the extended ECM/ADL model. When the data contain long memory, an ECM or ADL formulation would require a lot of lags of the variables to be included, a requirement which our parsimonious ACF correction disposes of. Using the matrix companion form, Abadir et al. (1999) showed that long lags have a similar effect to adding dimensions to a VAR (Vector AR), which would increases the bias and variance of the estimators, and is therefore detrimental to the modelling exercise.

Unlike traditional two-step GLS, we do not estimate the relation between $\boldsymbol{z}$ and $\boldsymbol{X}$ *then* correct for serial correlation or long memory. If the variables were not genuinely related, then their long memory would induce a spurious correlation problem, and the residuals from the regression of $\boldsymbol{z}$ on $\boldsymbol{X}$ should be of no value in correcting for the dynamics. Of course, the most efficient procedure would be to iterate between the two steps of GLS, until convergence obtains, or simply to use the ML procedure. However, if the simpler two-step procedure is adopted, then one should start by fitting the ACF. Note that, by conditioning in (14) on the explanatory variables $\boldsymbol{X}$, we have

$\boldsymbol{z} \sim \mathrm{D}_1(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, so the conditional ACFs of z_t and u_t are the same; see the comment following (4) and (13). Also, we are assuming that $\boldsymbol{X}$ is weakly exogenous (see Engle et al., 1983) for the parameters of D_1, which is justified in our two applications but need not be so in general. Otherwise, one needs joint modelling of $\boldsymbol{z}$ and $\boldsymbol{X}$, or estimation of the parameters of D_1 by means of instruments (effectively an orthogonal decomposition of the equation).

REFERENCES

Abadir, Karim M., Giovanni Caggiano and Gabriel Talmain, 2005, Nelson-Plosser revisited: the ACF approach, mimeo.

Abadir, Karim M., Kaddour Hadri and Elias Tzavalis, 1999, The influence of VAR dimensions on estimator biases, *Econometrica* 67, 163-181.

Abadir, Karim M. and Gabriel Talmain, 2002, Aggregation, persistence and volatility in a macro model, *Review of Economic Studies* 69, 749-779.

Abadir, Karim M. and A.M. Robert Taylor, 1999, On the definitions of (co-)integration, *Journal of Time Series Analysis* 20, 129-137.

Amemiya, Takeshi, 1985, *Advanced Econometrics* (Harvard University Press, Cambridge).

Baillie, Richard T., 1996, Long memory processes and fractional integration in econometrics, *Journal of Econometrics* 73, 5-60.

Baillie, Richard T. and Tim Bollerslev, 1994, Cointegration, fractional cointegration, and exchange rate dynamics, *Journal of Finance* 49, 737-745.

Baillie, Richard T. and Tim Bollerslev, 2000, The forward premium anomaly is not as bad as you think, *Journal of International Money and Finance* 19, 471-488.

Beran, Jan, 1992, Statistical methods for data with long-range dependence, *Statistical Science* 7, 404-427 (with discussion).

Cavaliere, Giuseppe, 2001, Testing the unit root hypothesis using generalized range statistics, *Econometrics Journal* 4, 70-88.

Chambers, Marcus J., 1998, Long memory and aggregation in macroeconomic time series, *International Economic Review* 39, 1053-1072.

Cooper, Michael J., Roberto C. Gutierrez and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345-1365.

De Bondt, Werner F.M. and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793-808.

De Bondt, Werner F.M. and Richard Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557-581.

Diebold, Francis X., Javier Gardeazabal and Kamil Yilmaz, 1994, On cointegration and exchange rate dynamics, *Journal of Finance* 49, 727-735.

Diebold, Francis X. and Glenn D. Rudebusch, 1989, Long memory and persistence in aggregate output, *Journal of Monetary Economics* 24, 189-209.

Engel, Charles, 1996, The forward discount anomaly and the risk premium: a survey of recent evidence, *Journal of Empirical Finance* 3, 123-192.

Engel, Charles and James D. Hamilton, 1990, Long swings in the dollar: are they in the data and do markets know it? *American Economic Review* 80, 689-713.

Engel, Charles and Kenneth D. West, 2004, Accounting for exchange rate variability in present-value models when the discount factor is near one, *American Economic Review* 94, 119-125.

Engle, Robert F., David F. Hendry and Jean-Francois Richard, 1983, Exogeneity, *Econometrica* 51, 277-304.

Fisher, Irving, 1930, *The Theory of Interest* (Macmillan, New York).

Gil-Alaña, Luis A. and Peter M. Robinson, 1997, Testing of unit root and other nonstationary hypotheses in macroeconomic time series, *Journal of Econometrics* 80, 241-268.

Griliches, Zvi, 1961, A note on serial correlation bias in estimates of distributed lags, *Econometrica* 29, 65-73.

Hendry, David F., Adrian R. Pagan and J. Denis Sargan, 1984, Dynamic specification, in Z. Griliches and M.D. Intriligator, eds.: *Handbook of Econometrics*, vol. 2 (North-Holland, Amsterdam).

Hendry, David F. and Thomas Von Ungern-Sternberg, 1980, Liquidity and inflation effects on consumers' expenditure, in A. S. Deaton (ed), *Essays in the Theory and Measurement of Consumers' Behaviour*. (Cambridge University Press, Cambridge). Reprinted In Hendry, D. F. (1993). *Econometrics: Alchemy or Science?* (Basil Blackwell, Oxford).

Lewis, Karen K., 1995, Puzzles in international financial markets, in G. Grossman and K. Rogoff, eds.: *Handbook of International Economics*, vol. 3 (North-Holland, Amsterdam).

Maddala, G.S. and A.S. Rao, 1973, Tests for serial correlation in regression models with lagged dependent variables and serially correlated errors, *Econometrica* 47, 761-774.

Nelson, Charles R. and Charles I. Plosser (1982) Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics* 10, 139-162.

Robinson, Peter M., 1994, Time series with strong dependence, in C.A. Sims, ed.: *Advances in Econometrics: sixth world congress*, vol.1 (Cambridge University Press, Cambridge).

Talmain, Gabriel, 2003, Stock market valuation and monopolistic competition: a dynamic stochastic general equilibrium approach, mimeo., Department of Economics, University of York.

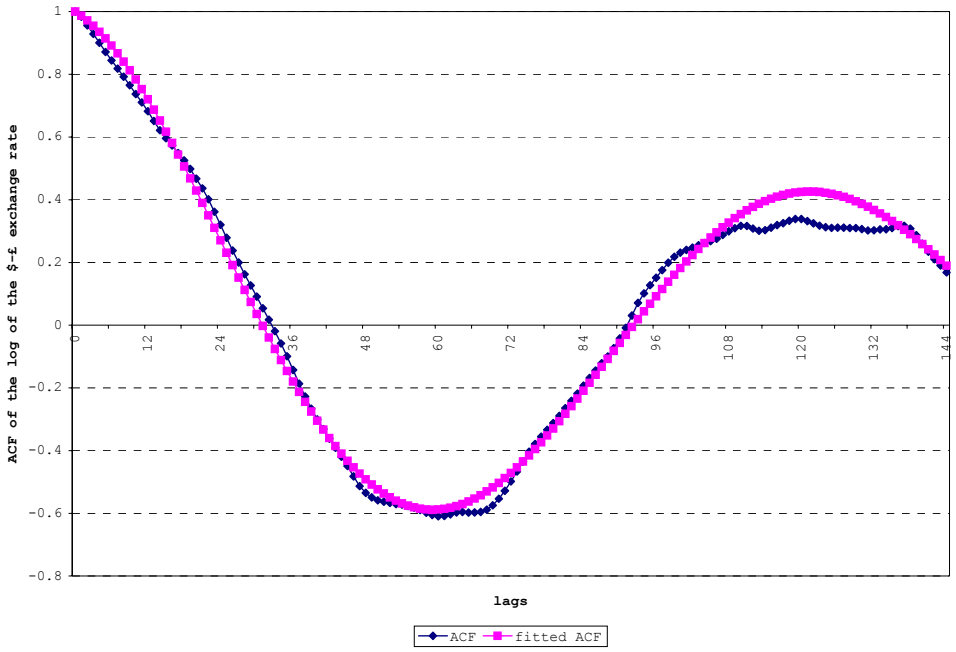

Figure 1. ACF of the logarithm of the exchange rate and its fit by our functional form.

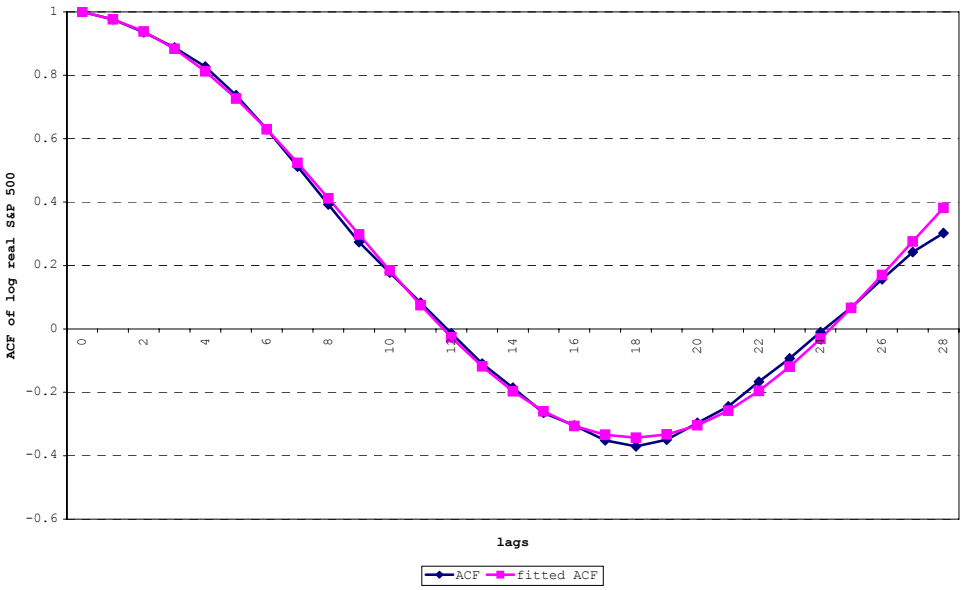

Figure 2. ACF of the logarithm of S&P 500 and its fit by our functional form.

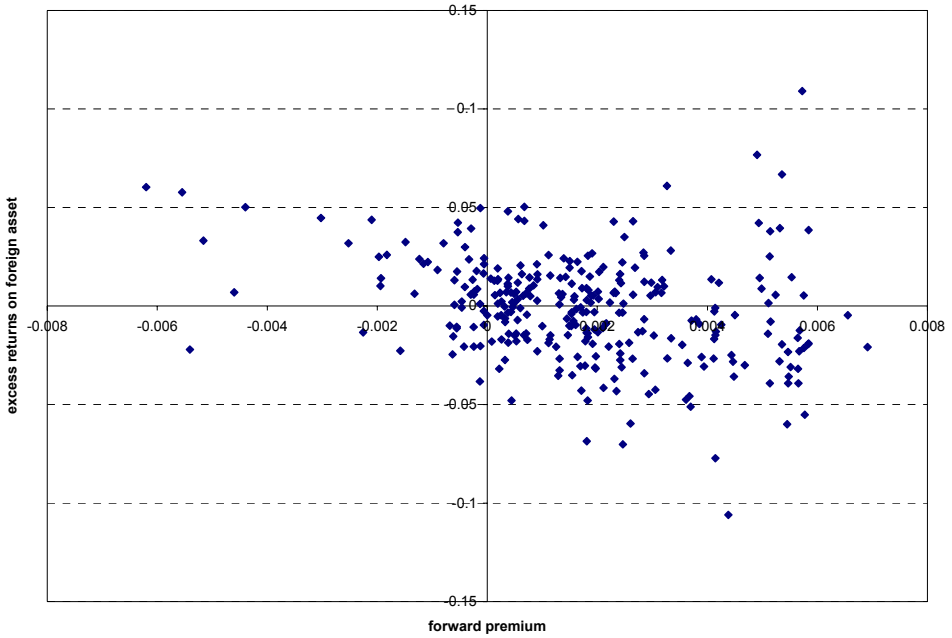

Figure 3. Scatter plot of excess returns vs. the forward premium, for the original data.

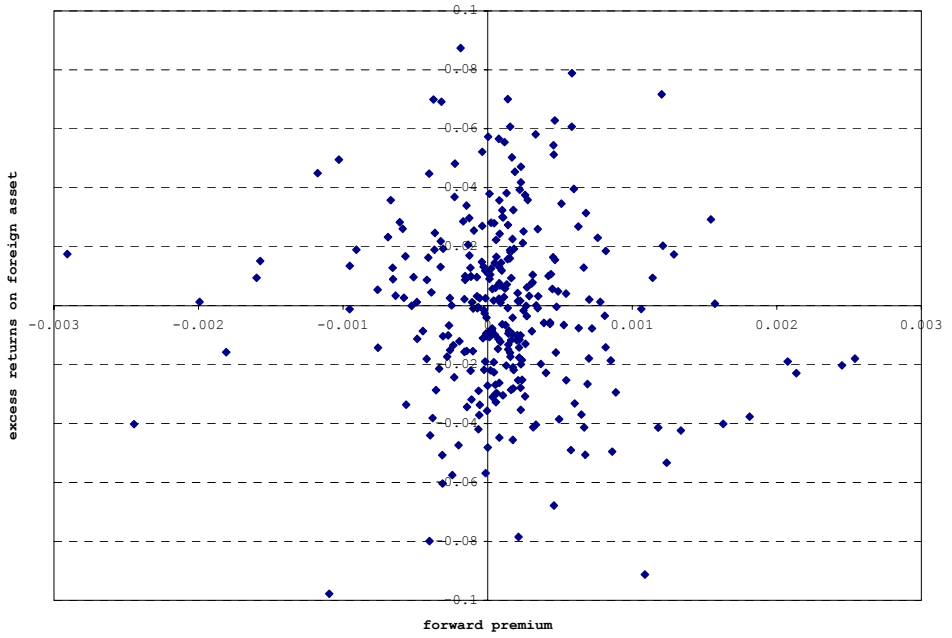

Figure 4. Scatter plot of excess returns vs. the forward premium, after ACF transformation.

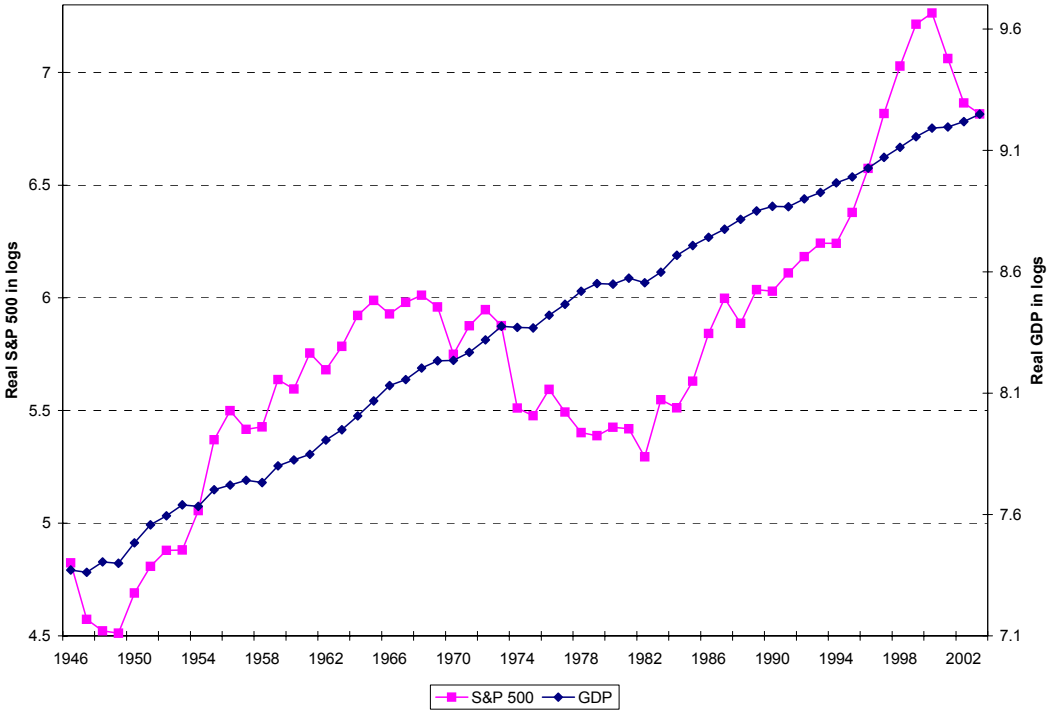

Figure 5. Logarithms of real S&P 500 and real GDP.

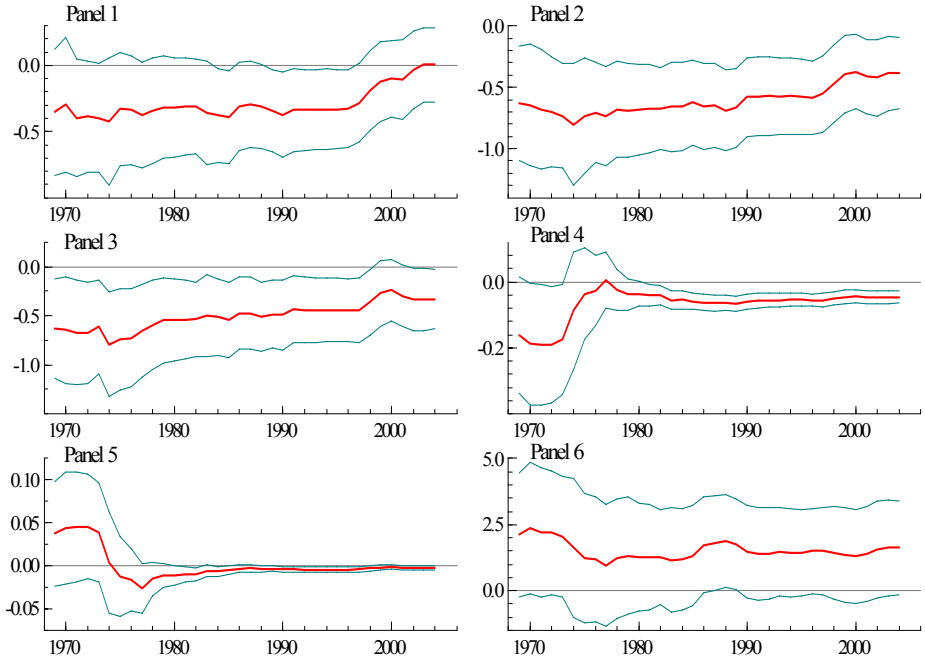

Figure 6. Recursive parameter estimates as the sample size is increased, with ±2 standard-error bands. Panels 1-6 represent the parameter estimates of Δs_{t-1}, Δs_{t-2}, Δs_{t-3}, x_{t-1}, t, Δy_{t-2}, respectively.

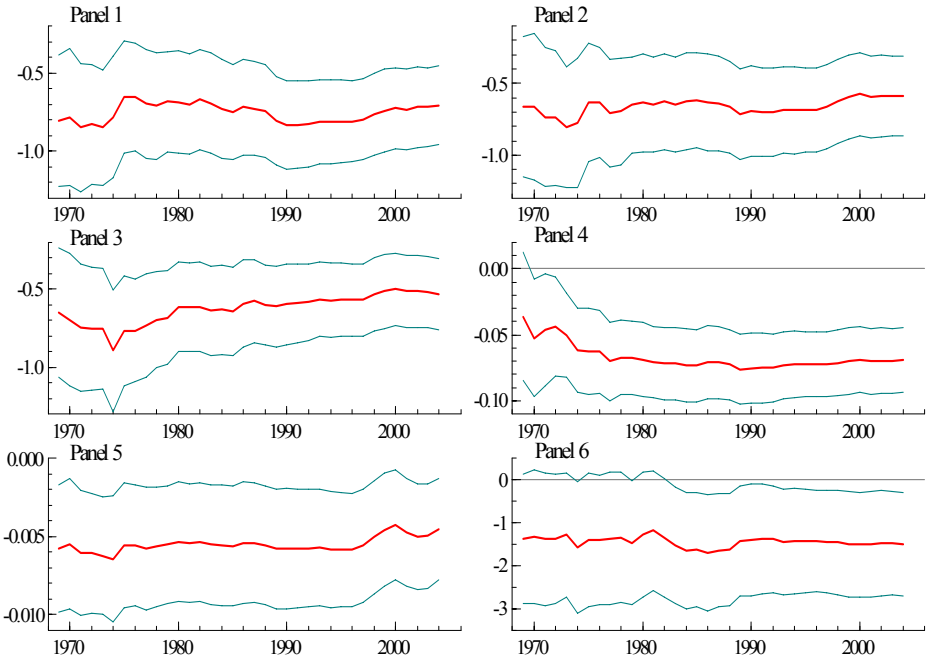

Figure 7. Recursive parameter estimates as the sample size is increased, with ±2 standard-error bands. Panels 1-6 represent the parameter estimates of Δs_{t-1}, Δs_{t-2}, Δs_{t-3}, x_{t-1}, t, Δy_{t-1}, respectively.

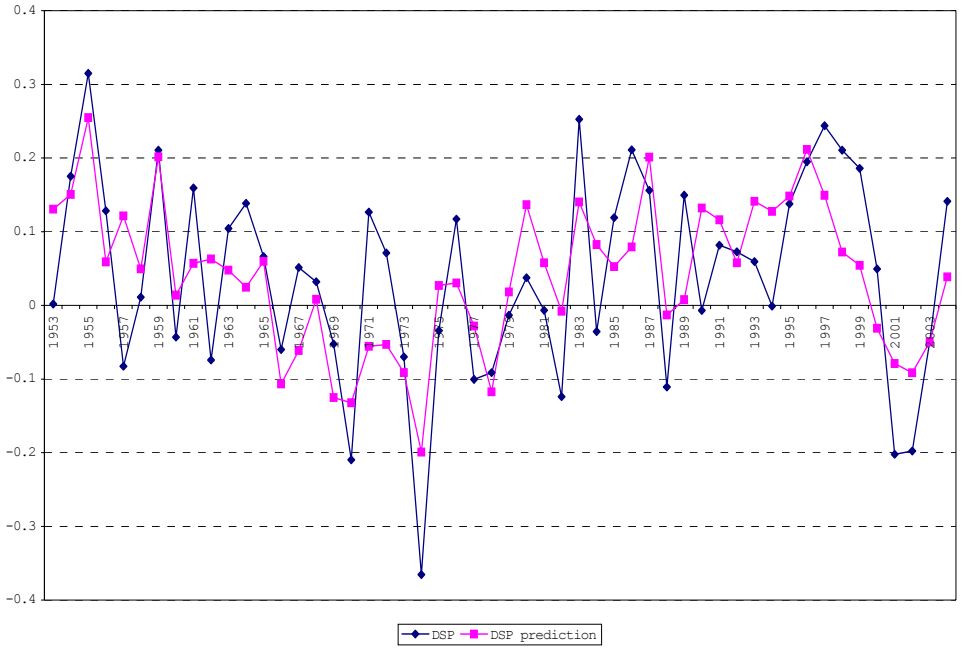

Figure 8. Actual and predicted changes in the logarithm of S&P 500.

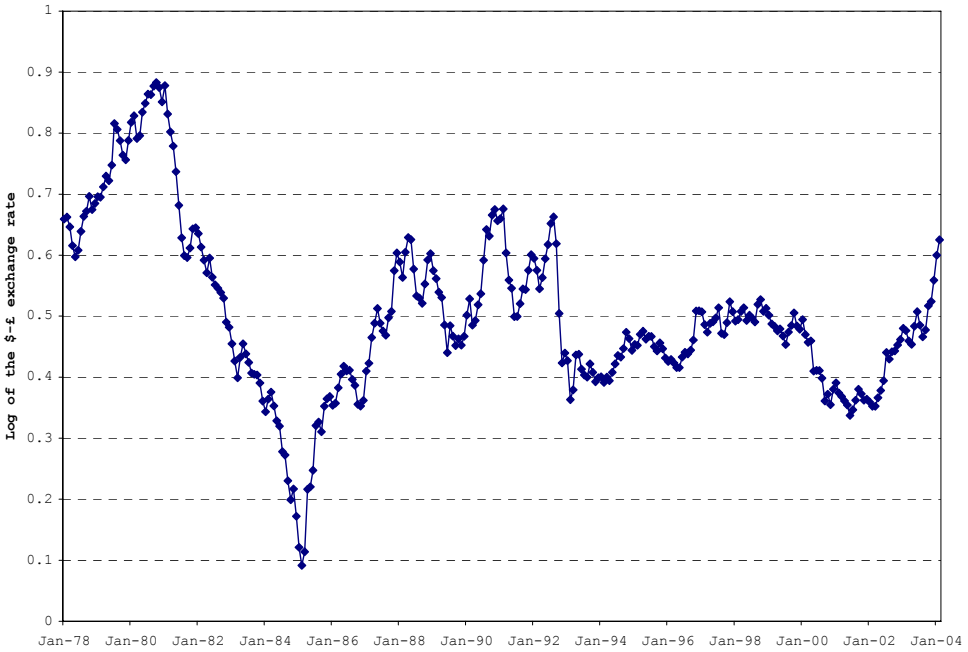

Figure 9. Logarithm of the $- £ exchange rate 1978-2004.

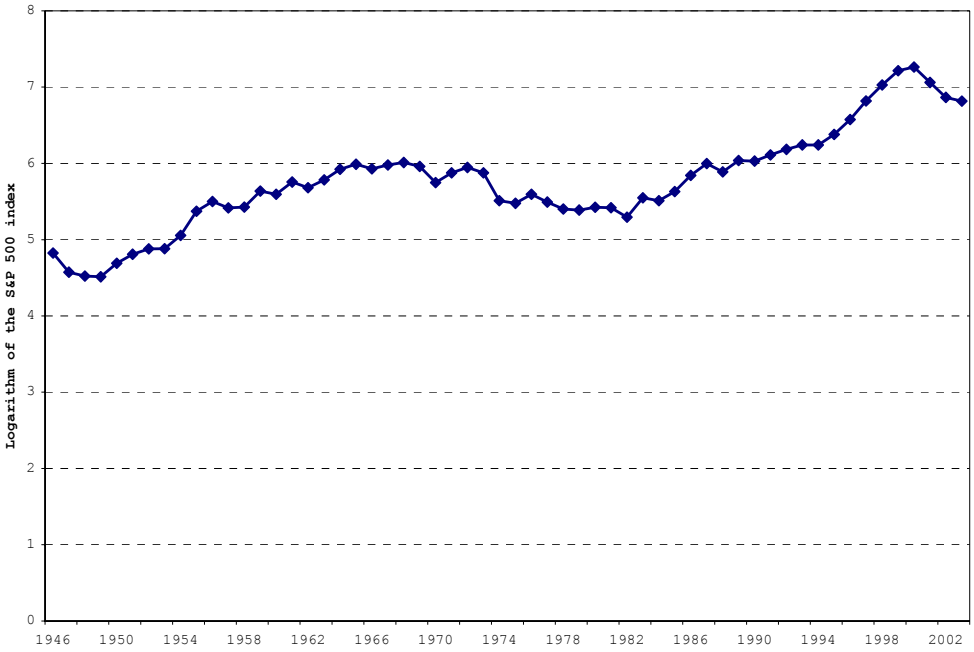

Figure 10. Logarithm of the S&P 500 stock index, 1946-2003.

European Central Bank working paper series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.int)

490 "Unions, wage setting and monetary policy uncertainty" by H. P. Grüner, B. Hayo and C. Hefeker, June 2005.

491 "On the fit and forecasting performance of New-Keynesian models" by M. Del Negro, F. Schorfheide, F. Smets and R. Wouters, June 2005.

492 "Experimental evidence on the persistence of output and inflation" by K. Adam, June 2005.

493 "Optimal research in financial markets with heterogeneous private information: a rational expectations model" by K. Tinn, June 2005.

494 "Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs" by A. Afonso and M. St. Aubyn, June 2005.

495 "Measuring inflation persistence: a structural time series approach" by M. Dossche and G. Everaert, June 2005.

496 "Estimates of the open economy New Keynesian Phillips curve for euro area countries" by F. Rumler, June 2005.

497 "Early-warning tools to forecast general government deficit in the euro area: the role of intra-annual fiscal indicators" by J. J. Pérez, June 2005.

498 "Financial integration and entrepreneurial activity: evidence from foreign bank entry in emerging markets" by M. Giannetti and S. Ongena, June 2005.

499 "A trend-cycle(-season) filter" by M. Mohr, July 2005.

500 "Fleshing out the monetary transmission mechanism: output composition and the role of financial frictions" by A. Meier and G. J. Müller, July 2005.

501 "Measuring comovements by regression quantiles" by L. Cappiello, B. Gérard, and S. Manganelli, July 2005.

502 "Fiscal and monetary rules for a currency union" by A. Ferrero, July 2005

503 "World trade and global integration in production processes: a re-assessment of import demand equations" by R. Barrell and S. Dées, July 2005.

504 "Monetary policy predictability in the euro area: an international comparison" by B.-R. Wilhelmsen and A. Zaghini, July 2005.

505 "Public good issues in TARGET: natural monopoly, scale economies, network effects and cost allocation" by W. Bolt and D. Humphrey, July 2005.

506 "Settlement finality as a public good in large-value payment systems" by H. Pagès and D. Humphrey, July 2005.

507 "Incorporating a "public good factor" into the pricing of large-value payment systems" by C. Holthausen and J.-C. Rochet, July 2005.

508 "Systemic risk in alternative payment system designs" by P. Galos and K. Soramäki, July 2005.

509 "Productivity shocks, budget deficits and the current account" by M. Bussière, M. Fratzscher and G. J. Müller, August 2005.

510 "Factor analysis in a New-Keynesian model" by A. Beyer, R. E. A. Farmer, J. Henry and M. Marcellino, August 2005.

511 "Time or state dependent price setting rules? Evidence from Portuguese micro data" by D. A. Dias, C. R. Marques and J. M. C. Santos Silva, August 2005.

512 "Counterfeiting and inflation" by C. Monnet, August 2005.

513 "Does government spending crowd in private consumption? Theory and empirical evidence for the euro area" by G. Coenen and R. Straub, August 2005.

514 "Gains from international monetary policy coordination: does it pay to be different?" by Z. Liu and E. Pappa, August 2005.

515 "An international analysis of earnings, stock prices and bond yields" by A. Durré and P. Giot, August 2005.

516 "The European Monetary Union as a commitment device for new EU member states" by F. Ravenna, August 2005.

517 "Credit ratings and the standardised approach to credit risk in Basel II" by P. Van Roy, August 2005.

518 "Term structure and the sluggishness of retail bank interest rates in euro area countries" by G. de Bondt, B. Mojon and N. Valla, September 2005.

519 "Non-Keynesian effects of fiscal contraction in new member states" by A. Rzońca and P. Ciżkowicz, September 2005.

520 "Delegated portfolio management: a survey of the theoretical literature" by L. Stracca, September 2005.

521 "Inflation persistence in structural macroeconomic models (RG10)" by R.-P. Berben, R. Mestre, T. Mitrakos, J. Morgan and N. G. Zonzilos, September 2005.

522 "Price setting behaviour in Spain: evidence from micro PPI data" by L. J. Álvarez, P. Burriel and I. Hernando, September 2005.

523 "How frequently do consumer prices change in Austria? Evidence from micro CPI data" by J. Baumgartner, E. Glatzer, F. Rumler and A. Stiglbauer, September 2005.

524 "Price setting in the euro area: some stylized facts from individual consumer price data" by E. Dhyne, L. J. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünnemann, F. Rumler and J. Vilmunen, September 2005.

525 "Distilling co-movements from persistent macro and financial series" by K. Abadir and G. Talmain, September 2005.