Caporale, Guglielmo Maria; Gil-Alana, Luis A.

Working Paper

Fractional integration and cointegration in US financial time series data

DIW Discussion Papers, No. 1116

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Caporale, Guglielmo Maria; Gil-Alana, Luis A. (2011) : Fractional integration and cointegration in US financial time series data, DIW Discussion Papers, No. 1116, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/55314

Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

Discussion Papers

1116

Guglielmo Maria Caporale • Luis A. Gil-Alana

Fractional Integration and Cointegration in US Financial Time Series Data

Berlin, March 2011

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Fractional Integration and Cointegration in US Financial Time Series Data

Guglielmo Maria Caporale[a]
Brunel University, London

Luis A. Gil-Alana[b]
University of Navarra

March 2011

Abstract

This paper examines several US monthly financial time series data using fractional integration and cointegration techniques. The univariate analysis based on fractional integration aims to determine whether the series are I(1) (in which case markets might be efficient) or alternatively I(d) with d < 1, which implies mean reversion. The multivariate framework exploiting recent developments in fractional cointegration allows to investigate in greater depth the relationships between financial series. We show that there exist many (fractionally) cointegrated bivariate relationships among the variables examined.

Keywords: Fractional integration, long-range dependence, fractional cointegration, financial data.

JEL Classification: C22, G10

[a] **Corresponding author:** Professor Guglielmo Maria Caporale, Research Professor at DIW Berlin. Centre for Empirical Finance, Brunel University, West London, UB8 3PH, UK. Tel.: +44 (0)1895 266713. Fax: +44 (0)1895 269770. *Email:* Guglielmo-Maria.Caporale@brunel.ac.uk

[b] The second-named author gratefully acknowledges financial support from the Ministerio de Ciencia y Tecnología (ECO2008-03035 ECON Y FINANZAS, Spain) and from a PIUNA Project of the University of Navarra.

1. Introduction

This paper re-examines the statistical properties of a number of US financial series (such as stock market prices, dividends, earnings, consumer prices, long-term interest rates) contained in the well-known dataset which can be downloaded from Robert Shiller's homepage, and which also are described in chapter 26 of Shiller's (1989) book on "Market Volatility".

In the existing literature, the Efficient Markets Hypothesis (EMH) has recently been tested using the present value (PV) model of stock prices, since, if stock market returns are not predictable, as implied by the EMH, stock prices should equal the present value of expected future dividends, and therefore stock prices and dividends should be cointegrated, as pointed out by Campbell and Shiller (1987). In their seminal paper, they tested the PV model of stock prices adopting Engle and Granger's (1987) cointegration procedure, an approach which is valid provided stock prices and dividends are stationary in first differences rather than in levels.[1] They used the Standard and Poor's (S&P's) dividends and value-weighted and equally-weighted New York Stock Exchange (NYSE) 1926-1986 datasets. In the case of the S&P series they rejected the unit root hypothesis for dividends but not for stock prices, whilst they could not reject it for either when using the NYSE data. As for cointegration, their results were also mixed, some test statistics rejecting the null hypothesis of no-cointegration, other failing to reject it. Han (1996) used Johansen's (1991) maximum likelihood (ML) method, and found that the deterministic cointegration restriction can be rejected on the basis of Canonical Cointegrating Regression (CCR) tests, and that stochastic cointegration is also rejected.[2]

[1] A constant discount rate is assumed in that study. In a subsequent paper (Campbell and Shiller, 1988) this assumption is relaxed to allow for time-varying discount rates in the PV model.

[2] Other empirical papers analysing cointegration in stock markets are Hakkio and Rush (1987), Baillie and Bollerslev (1989), Richards (1995), Crowder (1996), Rangvid (2001); other studies, such as Narayan

However, the discrete options I(1) and I(0) of classical cointegration analysis are rather restrictive: the equilibrium errors might in fact be a fractionally integrated I(d)-type process, with stock and dividends being fractionally cointegrated. This is stressed by Caporale and Gil-Alana (2004), who propose a simple two-step residuals-based strategy for fractional cointegration based on the approach of Robinson (1994a): first the order of integration of the individual series is tested, and then the degree of integration of the estimated residuals from the cointegrating regression. They find that the cointegrating relationship between stock prices and dividends possesses long memory, implying that the adjustment to equilibrium takes a long time and that PV models of stock prices are valid only over a long horizon.

The present study makes the following twofold contribution. Firstly, it applies univariate tests based on long memory techniques in order to establish the order of integration of the individual series and whether or not they are mean-reverting (which provides information about the empirical validity of the efficient market hypothesis). Therefore, compared to earlier studies, it extends the univariate analysis from the I(1)/I(0) cases to the more general case of fractional integration, which allows for a greater degree of flexibility in the dynamic specification. Secondly, it examines bivariate relationships among the variables using the most recent techniques in a fractional cointegration framework, which also allows for slow adjustment to equilibrium. To our knowledge, although numberless studies exist analysing such relationships, ours is the first to do so within such a framework. The implications of the findings are also discussed.

The layout of the paper is the following. Section 2 reviews the concepts of fractional integration and cointegration and the methods applied in this study. Section 3

---

and Smyth (2005) and Subramanian (2008) use instead cointegration techniques to analyse linkages between stock markets.

describes the data and reports the empirical results. Section 4 offers some concluding remarks.

2. Methodology

The methodology employed in this study is based on the concept of long memory or long range dependence. Given a zero-mean covariance stationary process $\{x_t, t = 0,\pm1,...\}$ with autocovariance function $\gamma_u = E(x_t, x_{t+u})$, in the time domain, long memory is defined such that:

$$\sum_{u=-\infty}^{\infty} |\gamma_u| = \infty.$$

Now, assuming that x_t has an absolutely continuous spectral distribution function, with a spectral density function given by:

$$f(\lambda) = \frac{1}{2\pi}\left(\gamma_0 + 2\sum_{u=1}^{\infty}\gamma_u \cos(\lambda u)\right),$$

according to the frequency domain definition of long memory the spectral density function is unbounded at some frequency λ in the interval $[0, \pi)$. Most of the empirical literature has focused on the case where the singularity or pole in the spectrum occurs at the 0-frequency. This is the standard case of I(d) models of the form:

$$(1 - L)^d x_t = u_t, \quad t = 0,\pm1,..., \qquad (1)$$

where L is the lag-operator ($Lx_t = x_{t-1}$) and u_t is I(0).[3] However, fractional integration may also occur at some other frequencies away from 0, as in the case of seasonal/cyclical models.

In the multivariate case, the natural extension of fractional integration is the concept of fractional cointegration. Though the original idea of cointegration, as in

[3] Throughout the paper we assume that $x_t = 0$ for $t \leq 0$. In other words, we adopt the Type I definition of fractional integration.

Engle and Granger (1987), allows for fractional orders of integration, all the empirical work carried out during the 1990s was restricted to the case of integer degrees of differencing. Only in recent years have fractional values also been considered. In what follows, we briefly describe the methodology used in this paper for testing fractional integration and cointegration in the case of Shiller's financial time series data.

2a. Fractional integration

There exist several methods for estimating and testing the fractional differencing parameter d. Some of them are parametric while others are semiparametric and can be specified in the time or in the frequency domain. In this paper, we use first a parametric approach developed by Robinson (1994a). This is a testing procedure based on the Lagrange Multiplier (LM) principle that uses the Whittle function in the frequency domain. It tests the null hypothesis:

$$H_o : \ d = d_o, \qquad (2)$$

for any real value d_o, in a model given by the equation (1), where x_t can be the errors in a regression model of the form:

$$y_t = \beta^T z_t + x_t, \qquad t = 1, 2, ..., \qquad (3)$$

where y_t is the observed time series, β is a (kx1) vector of unknown coefficients and z_t is a set of deterministic terms that might include an intercept (i.e., $z_t = 1$), an intercept with a linear time trend ($z_t = (1, t)^T$), or any other type of deterministic processes. Robinson (1994a) showed that, under certain very mild regularity conditions, the LM-based statistic $(\hat{r})$

$$\hat{r} \ \rightarrow_{dtb} \ N(0, 1) \qquad as \ \ n \rightarrow \infty,$$

where " $\rightarrow_{dtb}$ " stands for convergence in distribution, and this limit behaviour holds independently of the regressors used in (3) and the specific model for the I(0)

disturbances u_t in (1). The functional form of this procedure can be found in any of the numerous empirical applications based on his tests (see, e.g., Gil-Alana and Robinson, 1997; Gil-Alana and Henry, 2003; Cunado et al., 2005, etc.).

As in other standard large-sample testing situations, Wald and LR test statistics against fractional alternatives have the same null and limit theory as the LM test of Robinson (1994a). Lobato and Velasco (2007) essentially employed such a Wald testing procedure, and, although this and other recent methods such as the one developed by Demetrescu, Kuzin and Hassler (2008) have been shown to be robust with respect to even unconditional heteroscedasticity (Kew and Harris, 2009), they require an efficient estimate of d, and therefore the LM test of Robinson (1994a) seems computationally more attractive.[4]

In addition, we employ a semiparametric method (Robinson, 1995a) which is essentially a local 'Whittle estimator' in the frequency domain, using a band of frequencies that degenerates to zero. The estimator is implicitly defined by:

$$\hat{d} = \arg\min_d \left(\log \overline{C(d)} - 2\, d\, \frac{1}{m} \sum_{s=1}^{m} \log \lambda_s \right), \qquad (4)$$

$$\overline{C(d)} = \frac{1}{m} \sum_{s=1}^{m} I(\lambda_s)\, \lambda_s^{2d}, \qquad \lambda_s = \frac{2\pi s}{n}, \qquad \frac{1}{m} + \frac{m}{n} \to 0,$$

where $I(\lambda_s)$ is the periodogram of the raw time series, x_t, given by:

$$I(\lambda_s) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} x_t\, e^{i \lambda_s t} \right|^2,$$

and $d \in (-0.5, 0.5)$. Under finiteness of the fourth moment and other mild conditions, Robinson (1995a) proved that:

$$\sqrt{m}\, (\hat{d} - d^*) \to_{dtb} N(0, 1/4) \quad as \; n \to \infty,$$

[4] Other parametric estimation approaches (Sowell, 1992; Beran, 1995) were also employed for the empirical analysis producing very similar results as those obtained using the method of Robinson (1994a).

where d^* is the true value of d. This estimator is robust to a certain degree of conditional heteroscedasticity (Robinson and Henry, 1999) and is more efficient than other semi-parametric competitors.[5]

Although there exist further refinements of this procedure (Velasco, 1999, Velasco and Robinson, 2000; Phillips and Shimotsu, 2004; Shimotsu and Phillips, 2005; Abadir et al., 2007), these methods require additional user-chosen parameters, and the estimates of d may be very sensitive to the choice of these parameters. In this respect, the method of Robinson (1995a) seems computationally simpler and therefore preferable.

2b. Fractional cointegration

Engle and Granger (1987) suggested that, if two processes x_t and y_t are both I(d), then it is generally true that for a certain scalar $a \neq 0$, a linear combination $w_t = y_t - ax_t$, will also be I(d), although it is possible that w_t be I(d - b) with $b > 0$. This is the concept of cointegration, which they adapted from Granger (1981) and Granger and Weiss (1983). Given two real numbers d, b, the components of the vector z_t are said to be cointegrated of order d, b, denoted $z_t \sim CI(d, b)$ if:

(i) all the components of z_t are I(d),

(ii) there exists a vector $\alpha \neq 0$ such that $s_t = \alpha' z_t \sim I(\gamma) = I(d - b)$, $b > 0$.

Here, α and s_t are called the cointegrating vector and error respectively.[6] This prompts consideration of an extension of Phillips' (1991a) triangular system, which for a very simple bivariate case is:

[5] Other semiparametric univariate methods (e.g. the log-periodogram estimator of Robinson, 1995b) will be employed in the multivariate analysis based on fractional cointegration.

[6] Even considering only integer orders of integration, a more general definition of cointegration than the one given by Engle and Granger (1987) is possible, allowing for a multivariate process with components having different orders of integration. Nevertheless, in this paper we focus exclusively on bivariate cases and a necessary condition is that the two series display the same integration order.

$$y_t = \nu x_t + u_{1t}(-\gamma), \tag{5}$$

$$x_t = u_{2t}(-d), \tag{6}$$

for t = 0, ±1, ..., where for any vector or scalar sequence w_t, and any c, we introduce the notation $w_t(c) = (1 - L)^c w_t$. $u_t = (u_{1t}, u_{2t})^T$ is a bivariate zero mean covariance stationary I(0) unobservable process and $\nu \neq 0$, $\gamma < d$. Under (5) and (6), x_t is I(d), as is y_t by construction, while the cointegrating error $y_t - \nu x_t$ is I(γ). Model (5) and (6) reduces to the bivariate version of Phillips' (1991a) triangular form when γ = 0 and d = 1, which is one of the most popular models displaying CI(1, 1) cointegration considered in both the empirical and theoretical literature. Moreover, this model allows greater flexibility in representing equilibrium relationships between economic variables than the traditional CI(1, 1) prescription.

Next, we focus on the estimation of the cointegrating relationship, and in particular on the estimation of ν in (5) and (6). The simplest approach is to estimate it using the ordinary least squares (OLS) estimator

$$\hat{\nu}_{ols}{}^{t} = \frac{\sum_{t=1}^{n} x_t y_t}{\sum_{t=1}^{n} x_t^2}, \tag{7}$$

where the superscript "t" indicates time domain estimation. Here, in the standard cointegrating setting, with γ = 0 and d = 1, it has been shown (see, e.g., Phillips and Durlauf, 1986) that in general $\hat{\nu}_{ols}{}^{t}$ is n-consistent with non-standard asymptotic distribution. In fractional settings, the properties of OLS could be very different from those within this standard framework. When the observables are purely nonstationary (so that d ≥ 0.5), consistency of $\hat{\nu}_{ols}{}^{t}$ is retained, but its rate of convergence and

asymptotic distribution depends crucially on γ and d.[7] An alternative method of estimating ν is in the frequency domain. Consider the estimator

$$\hat{\nu}_{ols}{}^{f} = \frac{\sum_{j=0}^{n-1} I_{xy}(\lambda_j)}{\sum_{j=0}^{n-1} I_x(\lambda_j)}, \tag{8}$$

where $\lambda_j = 2\pi j/n$, j = 1, ..., n, are the Fourier frequencies, and for arbitrary sequences $\xi_t, \zeta_t,$ (possibly the same one as ξ_t), we define the discrete Fourier transform and (cross) -periodogram

$$w_\xi(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} \xi_t e^{it\lambda}, \qquad I_{\xi\zeta}(\lambda) = w_\xi(\lambda) w'_\zeta(-\lambda), \qquad I_\xi(\lambda) = I_{\xi\xi}(\lambda).$$

Here, the discrete Fourier transform at a given frequency captures the components of the series related to this particular frequency. Thus, noting that cointegration is a long-run phenomenon, when estimating ν one could concentrate just on low frequencies, which are precisely those representing the long-run components of the series, hence neglecting information from the high frequencies, associated with the short run, which could have a distorting effect on estimation. Robinson (1994b) proposed the narrow band least squares (NBLS) estimator, which is related to the band estimator proposed by Hannan (1963), and is given by

$$\hat{\nu}_{NBLS} = \frac{\sum_{j=0}^{m} s_j \operatorname{Re} I_{xy}(\lambda_j)}{\sum_{j=0}^{m} s_j I_x(\lambda_j)}, \tag{9}$$

where $1 \leq m \leq n/2$, $s_j = 1$ for j = 0, n/2, 2, otherwise, and $(1/m) + (m/n) \to 0$ as $n \to \infty$. Robinson (1994b) showed the consistency of this estimator even under stationary cointegration, using the fact that focusing on a degenerating band of low frequencies

[7] Robinson (1994b) showed that under stationary cointegration (i.e. d < 0.5) the OLS estimator is inconsistent.

reduces the bias due to the contemporaneous correlation between u_{1t} and u_{2t}, which was precisely the reason why OLS was inconsistent in some cases. As with OLS, in general NBLS has a non-standard limiting distribution.

With the aim of obtaining estimates of ν with improved asymptotic properties (optimal rate of convergence, median unbiasedness, asymptotic mixed-normality leading to standard inference procedures), more refined techniques to estimate ν have been proposed in a fractional setting. These are related to the work of Johansen (1988, 1991), Phillips and Hansen (1990), Phillips (1991a,b), Phillips and Loretan (1991), Saikkonen (1991), Park (1992), and Stock and Watson (1993), who all proposed estimators with optimal asymptotic properties (under Gaussianity) in the standard cointegrating setting with γ = 0 and d = 1. However, for all these estimators knowledge of γ and d was assumed (usually after pretesting), which in fractional circumstances might be hard to justify.[8]

Assuming that the process u_t in (5) and (6) has a parametric spectral density $f(\lambda) = f(\lambda;\theta)$, where θ is an unknown vector of short-memory parameters, Robinson and Hualde (2003), based on generalized least squares (GLS)-type corrections, propose methods to estimate optimally (under Gaussianity) ν when d – γ > 0.5 (named strong cointegration). Denoting

$$z_t(c,d) = (y_t(c), x_t(d))', \ \zeta = (1,0)', \ p(\lambda;h) = \zeta' f(\lambda;h)^{-1},$$

$$a(c,d,h) = \sum_{j=1}^{n} p(\lambda_j;h) w_{x(c)}(-\lambda_j) w_{z(c,d)}(\lambda_j), \quad q(\lambda;h) = \zeta' f(\lambda;h)^{-1} \zeta,$$

$$b(c,d) = \sum_{j=1}^{n} q(\lambda_j;h) I_{x(c)}(\lambda_j), \text{ and defining } \hat{\nu}(c,d,h) = \frac{a(c,d,h)}{b(c,h)},$$

they considered five different estimators given by:

[8] Dolado and Marmol (1996) proposed an extension to the fractional setting of the Fully Modified (FM)-OLS estimator of Phillips and Hansen (1990), assuming knowledge of γ and d.

$$\hat{\nu}(\gamma,\delta,\theta),\ \hat{\nu}(\gamma,\delta,\hat{\theta}),\ \hat{\nu}(\gamma,\hat{\delta},\hat{\theta}),\ \hat{\nu}(\hat{\gamma},\delta,\hat{\theta}),\ \hat{\nu}(\hat{\gamma},\hat{\delta},\hat{\theta}), \qquad (10)$$

where $\hat{\gamma}, \hat{\delta}, \hat{\theta},$ are corresponding estimators of the nuisance parameters γ, d and θ. The estimators in (10) reflect different knowledge about the structure of the model, the first being in general unfeasible, the second only assuming knowledge of the integration orders (as was done previously in the standard cointegrating literature), whereas the last estimator represents the most realistic case. Under regularity conditions, Robinson and Hualde (2003) showed that any of the estimators in (10) is $n^{d-\gamma}$-consistent with identical mixed-Gaussian asymptotic distributions, leading to Wald tests on the parameter ν,

$$W(\gamma,\delta,\theta),\ W(\gamma,\delta,\hat{\theta}),\ W(\gamma,\hat{\delta},\hat{\theta}),\ W(\hat{\gamma},\delta,\hat{\theta}),\ W(\hat{\gamma},\hat{\delta},\hat{\theta}), \qquad (11)$$

where $W(c,d,h) = b(c,h)\{\hat{\nu}(c,d,h) - 1\}^2$, with a chi-squared limit distribution.[9]

3. Data and Empirical results

The monthly series analysed have been collected by Robert Shiller and his associates, and are available on http://www.econ.yale.edu/~shiller/. The sample period goes from 1871m1 to 2010m6. They are described in chapter 26 of Shiller's (1989) book on "Market Volatility", where further details can be found, and are constantly updated and revised. Specifically, they are the following series: stock market prices (monthly averages of daily closing S&P prices, computed from the S&P four-quarter tools for the quarter since 1926, with linear interpolation to monthly figures); dividends (an index), earnings (also an index), a consumer price index (Consumer Price Index - All Urban Consumers) used for computing real values of the previous variables, a long-term

[9] Hualde and Robinson (2007) propose an estimator of ν in (5) and (6) in the case when $d - \gamma < 0.5$ (named weak cointegration). As in Robinson and Hualde (2003), this method is based on a GLS-type correction. Hualde and Robinson (2007) showed that the estimators are $n^{1/2}$-consistent and asymptotically normally distributed.

interest rate (GS10, which is the yield on the 10-year Treasury bonds), and also a cyclically adjusted price earnings ratio.

3a. Univariate analysis: fractional integration

We first employ the parametric approach of Robinson (1994a) described in Section 2, assuming that the disturbances are white noise. Thus, time dependence is exclusively modelled through the fractional differencing parameter d. In particular, we consider the set-up in (3) and (1), with $z^T = (1,t)^T$, testing H_o (2) for d_o-values equal to 0, (0.001), 2. In other words, the model under the null becomes:

$$y_t = \beta_0 + \beta_1 t + x_t; \qquad (1-L)^{d_o} x_t = u_t \qquad t = 1, 2, ...,$$

and white noise u_t.

Table 1 displays the estimates of d (obtained as the values of d_o that produce the lowest $\hat{r}-$statistics in absolute value)[10] along with the 95% confidence band of the non-rejection values of d using Robinson's (1994a) parametric approach. For each series, we display the three cases commonly examined in the literature, i.e., the cases of no regressors (i.e, $\beta_0 = \beta_1 = 0$), an intercept ($\beta_1 = 0$), and an intercept with a linear time trend.

[Insert Table 1 about here]

The first noticeable feature is that all the estimated values of d are above 1 and the unit root null hypothesis (i.e., d = 1) is rejected in all cases at the 5% level. In general the values are very similar for the three cases with deterministic terms, although the results change substantially from one series to another. Specifically, values of d above 1.5 are found in the case of dividends, earnings and real earnings. For the

[10] This is standard practice in the context of Robinson's (1994a) tests, and produces estimates that are very similar to the Whittle estimates in the frequency domain (Dahlhaus, 1989). Very similar values were obtained with other methods (Sowell, 1992; Beran, 1995).

remaining series the values are slightly above 1, but still significantly different from 1. However, these results might be biased due to the lack of (weak)-autocorrelation for the error term. Therefore, in what follows we assume that the disturbances are weakly autocorrelated and model them first using the exponential spectral model of Bloomfield (1973). This is a non-parametric approach to modelling the I(0) error term that produces autocorrelations decaying exponentially as in the AR(MA) case.[11] The results using this approach are displayed in Table 2.

[Insert Tables 2 and 3 about here]

It can be seen that the values are much smaller than in the previous case of white noise disturbances. One series (long-term interest rates) has values which are strictly below 1, implying mean-reverting behaviour; for dividends and real stock prices the unit root null cannot be rejected. It is slightly rejected (at the 5% level but not at the 10% level) for stock prices, consumer price index and price/earning ratio, and it is decisively rejected in favour of higher orders of integration for the remaining two series (earnings and real earnings). As a final specification we assume that the error term follows a seasonal AR(1) process. The results (displayed in Table 3) are very similar to those based on white noise disturbances, with estimates of d which are all strictly above 1. Deeper inspection indicates that time trends are not required in any case, the intercept being sufficient for the deterministic component. Moreover, LR tests and other residuals-based tests suggest that the d-differenced series may all be weakly (non-seasonally) autocorrelated, implying that the model with Bloomfield disturbances may approximate accurately the order of integration of the series. Nevertheless, in view of the sensitiveness of the results to the specification of the error term, we also apply a

[11] This model is extremely well suited to Robinson's (1994a) tests (see Gil-Alana, 2004).

semiparametric method that does not specify a functional form for the I(0) disturbance term.

[Insert Figure 1 and Table 4 about here]

Figure 1 displays for each series the estimates of d based on the semiparametric method of Robinson (1995a), i.e., $\hat{d}$ as given by (4). The estimates were obtained based on the first differenced series, then adding 1 to get the proper estimates of d. The estimates of d are shown for a whole range of values of the bandwidth parameter m = 1, 2, …, n/2 (on the horizontal axis)[12]; the 95% confidence bands corresponding to the I(1) hypothesis are also displayed. It can be seen that, for small values of m, the unit root null is rejected in favour of mean reversion (d < 1) in the case of earnings, real dividends, real earning and price earning ratio. For the remaining series (still with a small m) the estimated values of d are within the I(1) interval, except for the CPI series for which d is found to be strictly above 1. However, when the bandwidth parameter is large, the estimates are clearly above 1 in all cases, the only exception being long-term interest rates, with many values in the I(1) interval. Table 4 reports the numerical values for different bandwidth parameters, m = 25, 41 (= $T^{0.5}$), 100, 200, 300 and 500: at the 95% level, there are several cases where the unit root null cannot be rejected.[13]

3b. Multivariate analysis: fractional cointegration

A number of cointegrating relationships might exist between the individual variables examined in the previous subsection, in particular between:

a) Stock prices and dividends

[12] The choice of the bandwidth is crucial in view of the trade-off between bias and variance: the asymptotic variance is decreasing with m while the bias is growing with m.

[13] Specifically, in the case of m = $(T)^{0.5}$, which has been widely considered in the empirical literature, the unit root null hypothesis cannot be rejected for stock prices, dividends, long-term interest rates, real stock prices and real dividends, whilst it is rejected in favour of mean reversion (i.e., d < 1) for earnings and real earnings, and in favour of d > 1 for the consumer price index.

b) Real stock prices and real dividends

c) Price/Earning ratio and long-term interest rates

and

d) Real stock prices and real earnings

Some of these relationships have been extensively analysed in the literature. Campbell and Shiller (1987) and DeJong (1992) tested a present value model of the stock market using time series data for real US annual prices and dividends from 1871 to 1986. In the first of these studies, they carried out ADF tests, with and without a time trend, on both individual series, and their results suggested that both series were integrated of order 1. When using the DF, ADF tests on the residuals from the cointegrating regressions, their results were mixed: the former test rejected the null hypothesis of no cointegration at the 5% level, while the latter narrowly failed to reject it at the 10% level. DeJong (1992) used a Bayesian approach to model these two variables and found evidence in favour of trend-stationary representations. Similarly, Koop (1991), using a different dataset, came to the same conclusion that both variables are stationary around a linear trend, and, even when assuming unit roots, he found little evidence of cointegration.

Pereira-Garmendia (2010) finds that real stock price and real earnings are related through inflation. The relationship between stock prices, earnings and bond yield is analysed by Durre and Giot (2007). Papers examining long-run linkages between the price/earnings ratio and interest rates include Phillips (1999), Campbell and Shiller (1998, 2001), and Asness (2003) inter alia.

In all cases, we follow the same strategy. We first estimate individually the orders of integration using now the log-periodogram-type estimator devised by Robinson (1995b). This is defined as:

$$\hat{d}(l) = \sum_{j=l+1}^{m} (a_j - \bar{a}) \log I(\lambda_j) / S_l, \qquad (12)$$

where

$$a_j = -\log\left(4\sin^2\left(\frac{\lambda_j}{2}\right)\right), \quad \bar{a} = \frac{1}{m-l}\sum_{j=1}^{p} a_j u_{t-j} + \varepsilon_t, \quad S_l = \sum_{j=l+1}^{m} (a_j - \bar{a})^2, \quad \lambda_j = \frac{2\pi j}{n},$$

and $0 \leq l < m < n$. The results for the individual series possibly involved in cointegration relationships are displayed in Table 5 (for $m = T^{0.5}$ and $l = 0, 1, \ldots, 5$).[14]

Next we test the homogeneity of the orders of integration in the bivariate systems (i.e., H_o: $d_x = d_y$), where d_x and d_y are the orders of integration of the two individual series, by using an adaptation of Robinson and Yajima (2002) statistic $\hat{T}_{xy}$ to log-periodogram estimation. The statistic is:

$$\hat{T}_{xy} = \frac{m^{1/2}(\hat{d}_x - \hat{d}_y)}{\left(\frac{1}{2}(1 - \hat{G}_{xy} / (\hat{G}_{xx}\hat{G}_{yy})\right)^{1/2} + h(n)}, \qquad (13)$$

where $h(n) > 0$ and $\hat{G}_{xy}$ is the $(xy)^{th}$ element of

$$\hat{G} = \frac{1}{m}\sum_{j=1}^{m} \mathrm{Re}\left[A\hat{\Lambda}(\lambda_j)^{-1} I(\lambda_j)\hat{\Lambda}(\lambda_j)^{-1*}\right], \quad \hat{\Lambda}(\lambda_j) = diag\left\{e^{i\pi\hat{d}_x/2}\lambda^{-\hat{d}_x}, e^{i\pi\hat{d}_y/2}\lambda^{-\hat{d}_y}\right\}.$$

The results using this approach are displayed in Table 6. In general, we cannot reject the null hypothesis of equal orders of integration.[15] In the following step, we perform the Hausman test for no cointegration of Marinucci and Robinson (2001) comparing the estimate $\hat{d}_x$ of d_x with the more efficient bivariate one of Robinson (1995b), which uses the information that $d_x = d_y = d_*$. Marinucci and Robinson (2001) show that

[14] We will examine later these tables in detail for each of the potential cointegrating relationships.
[15] As in the case of the previous table, the comments for the specific series will be presented later.

$$H_{is} = 8s\left(\hat{d}_* - \hat{d}_i\right)^2 \rightarrow_{dtb} \chi_1^2 \qquad as \quad \frac{1}{s} + \frac{s}{n} \rightarrow 0, \tag{14}$$

with i = x, y, and where s < [n/2] is a bandwidth parameter, analogous to m introduced earlier; $\hat{d}_i$ are univariate estimates of the parent series, and $\hat{d}_*$ is a restricted estímate obtained in the bivariate context under the assumption that $d_x = d_y$. In particular,

$$\hat{d}_* = -\frac{\sum_{j=1}^{s} 1_2' \hat{\Omega}^{-1} Y_j v_j}{2 1_2' \hat{\Omega}^{-1} 1_2 \sum_{j=1}^{s} v_j^2}, \tag{15}$$

with $Y_j = [\log I_{xx}(\lambda_j), \log I_{yy}(\lambda_j)]^T$, and $v_j = \log j - \frac{1}{s}\sum_{j=1}^{s} \log j$. The limiting distribution above is presented heuristically, but the authors argue that it seems sufficiently convincing for the test to warrant serious considerations. The results using this approach are displayed in Table 7.

In the final part of the analysis, we apply the methods of Robinson and Hualde (2003) and Hualde and Robinson (2007). We identify parametric models for f(λ) with u_t in (5) and (6) having the form,

$$u_t = A(L)\varepsilon_t, \tag{16}$$

where ε_t is supposed to be an i.i.d. process, and A(L) is diagonal, treating thus u_{1t} and u_{2t} separately. We approximate the two series as

$$u_{1t} = (1 - L)^{\tilde{\gamma}}\left[y_t - \hat{v}_{ols} x_t\right], \tag{17}$$

and

$$u_{2t} = (1 - L)^{\tilde{d}}\, x_t, \tag{18}$$

to obtain estimates of γ and d previously estimated using other methods, and follow Box-Jenkins-type procedures to identify the models within the ARMA class. The results based on this method are displayed in Tables 13a – 13d.

[Insert Tables 5, 6 and 7 about here]

Next we examine the bivariate relationships.

3.2.a Stock market prices and dividends

[Insert Figure 2a about here]

Figure 2a displays the plots of the two series. Both of them are relatively stable until the end of World War II, when they start increasing and also exhibit a higher degree of volatility.

Focusing first on the univariate results using the Whittle semiparametric estimator (Robinson, 1995a), it can be seen that for small values of m the unit root null cannot be rejected (see Table 4). Specifically, for m = $(T)^{0.5}$ = 41, the estimates are 0.953 and 1.105 respectively for stock prices and dividends. Similar evidence of unit roots, though with slightly higher values, is obtained with the log-periodogram estimator of Robinson (1995b) (see Table 5). For example, for l = 0, 1, 2, …, 5, and m = $(T)^{0.5}$, the estimates of d for stock prices ranges between 1.041 and 1.080 and those for dividends between 1.026 and 1.222. Testing now the homogeneity condition with Robinson and Yajima's (2002) procedure (see Table 6), it is found that the two orders of integration are equal.[16] The Hausmann test of no cointegration (Marinucci and Robinson, 2001) (see Table 7) indicates that the estimates of d for the individual series using the bivariate representation ($\hat{d}_*$ in (15)) are very close to 1 and not significantly different from 1 (using three different values for s), and evidence of cointegration is only obtained in one case out of the six considered.

[16] Here h(n) is set equal to b^{-5-2i}, with i = 1, 2, 3, 4 and 5 and b = $(T)^{0.5}$, which is the bandwidth used in the estimation.

3.2.b Real stock market prices and real dividends

[Insert Figure 2b about here]

The same relationship as above but in real terms is examined in this subsection. A time series plot of the two series is displayed in Figure 2b. They exhibit a similar pattern to the previous case although with more volatility in the early part of the sample, and may have a common stochastic trend. Starting again with the univariate tests (see Table 4), it is found that, when applying the Whittle semiparametric method of Robinson (1995a), for $m = (T)^{0.5} = 41$, the estimates of d are 0.888 and 0.896 respectively for real stock prices and real dividends, and the unit root null cannot be rejected for either series. Similar evidence is obtained with the log-periodogram estimator (see Table 5), with values of d ranging from 0.972 and 1.085 for real stock prices, and from 0.822 and 0.997 for real dividends. The test of homogeneity of the orders of integration (Table 6) implies equality in the values of d, whilst testing the null of no cointegration with the Hausman test of Robinson and Marinucci (2001) suggests that the two series might be cointegrated.

3.2.c Price / earning ratio and long-term interest rates

[Insert Figure 2c about here]

These two series are plotted in Figure 2c. Interest rates appear to be more stable than the price/earning ratio during the first half of the sample; however, during the second half, there is a sharp increase in interest rates but not in the price/earning ratio. As for the Whittle estimates of d (see Table 4), it is found that for the price/earning ratio the values of d are very sensitive to the bandwidth parameter: for small values (e.g., 25, 41 or 100) the unit root is rejected in favour of values of d below 1; on the contrary, the unit root

null cannot be rejected for m = 200, and it is rejected in favour of d > 1 for m = 300 and 500. For the long-term interest rates, the results are more stable and the unit root null cannot be rejected for any bandwidth parameter. These results are corroborated by the log-periodogram estimates, displayed in Table 5. Thus, for the price/earning ratio, different results are obtained depending on whether or not the series is first-differenced, while for long-term interest rates the evidence strongly support the I(1) case. Interestingly, when performing the homogeneity tests of Robinson and Yajima (2002) we cannot reject the null of equal orders of integration, and the Hausman test reject in all cases the null hypothesis of no cointegration.

3.2.d Real stock market prices and real earnings

[Insert Figure 2d about here]

Plots of the two series are displayed in Figure 2d. They both have a very similar upward trend, which suggests that they may be cointegrated. The estimated values of d using the Whittle method and for $m = (T)^{0.5}$ (see Table 4) are 1.071 for real stocks and 0.933 for real earnings, and in both cases we cannot reject the null of I(1) series. The same evidence in favour of unit roots is obtained with the log-periodogram estimates in Table 5, and the homogeneity restriction cannot be rejected (see Table 6). The Hausman tests of Robinson and Marinucci (2001) also indicate that the two series might be cointegrated.

[Insert Table 8 about here]

Table 8 displays different estimates of the cointegrating coefficients for each of the four relations examined. These are found to be relatively stable across the different procedures.

Finally, on the basis of the coefficients displayed in Table 8, we estimated the orders of integration in the residuals of the cointegrating regression. First, we used the parametric approach of Robinson (1994a). However, the results vary considerably depending on the specification of the error term. Due to this disparity, we estimate d with semiparametric methods.

[Insert Tables 9 - 12 about here]

Tables 9 and 10 display the estimates of d based on the log periodogram regression estimator of Robinson (1995b) for $u = T^{0.5}$ and $l = 0$ and $l = 2$ respectively. In many cases the estimates are strictly smaller than 1, especially for the price/earning ratio – long-term interest rates and real stock prices – real earning relationships.

Tables 11 and 12 report the results from the semiparametric Whittle method of Robinson (1995a), again applied to the estimated residuals from the cointegrating relationships. Two different bandwidth parameters, $m = 25$ (in Table 11) and $m = T^{0.5} = 41$, are considered in Table 12. Virtually all estimated values are strictly below 1. For the first two relationships (stock prices and dividends and their real terms) the values for the order of integration in the residuals range between 0.6 and 0.8. Smaller values are obtained for the price/earning ratio – long-term interest rate relationship: if $m = 41$, the estimated value of d is about 0.55, however using $m = 25$, the values are in all cases 0.50 suggesting that the series may be stationary.[17] There is a wider range of values in the case of the real stock prices – real earnings relationship, although most of them are also in the interval (0.5, 1).

Finally, we identify parametric models for $f(\lambda)$ with u_t in (5) and (6) on the basis of equations (16) – (18), using wide-ranging values for the orders of integration from the previous tables. Using a Box-Jenkins-type methodology we identified at most AR(1)

[17] Note that these estimates are based on the first differenced data, and a value of 1 is then added to obtain the proper estimates of d.

structures in all cases. Therefore, we simply consider combinations of white noises and AR(1) processes in each bivariate relation. For each model, we apply the univariate Whittle procedure of Velasco and Robinson (2000), using untapered versions, and, as usual, the first-differenced data, then adding 1 to the estimated value. The results for the four bivariate relationships are summarised in Tables 13a – 13d and they are fairly similar for the different types of I(0) errors.

[Insert Table 13 about here]

Although we do not report it, we also estimated a multivariate version of the Bloomfield (1973) model for I(0) autocorrelation, with fairly similar results to those presented in Table 13. In general, there is a reduction in the order of integration of about 0.3/0.4 from the original series to the cointegrating relationship. The orders of integration in the latter are about 0.7 for three of these relations: stock prices/dividends; real prices/real dividends, and real prices/real earnings. For the price-earning ratio/interest rates relationship, the reduction is slightly bigger, and the order of integration of the cointegrating relationship seems to be slightly above 0.5.

4. Conclusions

In this paper we have examined bivariate relationships among various financial variables using some recent techniques based on the concepts of fractional integration and cointegration. In particular, we focus on the following bivariate relationships: stock prices and dividends; real stock prices and real dividends; price/earning ratio and long run interest rates, and real stock prices and real earnings, monthly, for the time period 1871m1 to 2010m6.

The univariate results strongly support the hypothesis that all individual series are nonstationary with orders of integration equal to or higher than 1 in practically all

cases. In fact, mean reversion is not found for any of the series examined.[18] The multivariate results indicate that the four bivariate relationships are fractionally cointegrated with the orders of integration of the cointegrating regressions being in the interval [0.5, 1) and therefore displaying mean reverting behaviour. The implication is that there exist long-run equilibrium relationships consistent with economic theory and that the effects of shocks are temporary, although the fact that fractional cointegration (rather than standard cointegration) holds means that the adjustment process is much slower, and that therefore the overall costs of deviations from equilibrium are bigger than standard cointegration approaches would estimate. This is an important result that should be taken into account when formulating policies and deciding on policy actions.

Admittedly, our analysis does not take into account other possible features of the data, such as structural breaks, non-linearities and other issues. Of course, these are also important issues whose relevance for fractional integration tests has already been investigated (see, e.g., Diebold and Inoue, 2001; Granger and Hyung, 2004; Caporale and Gil-Alana, 2008). Our future research will consider them in the context of fractional cointegration.

[18] A small degree of mean reversion is found in the long-term interest rates when using the parametric method of Robinson (1994a) with Bloomfield (1973)-type disturbances.

References

Asness, C.S., 2003, Fight the Fed Model. The Journal of Portfolio Management, vol. 30, 11-24.

Abadir, K.M., W. Distaso and L. Giraitis, 2007, Nonstationarity-extended local Whittle estimation, Journal of Econometrics 141, 1353-1384.

Baillie, R. and T. Bollerslev (1989), "Common stochastic trends in a system of exchange rates", Journal of Finance 44, 167-181.

Beran, J., 1995, Maximum likelihood estimation of the differencing parameter for invertible short and long memory ARIMA models, Journal of the Royal Statistical Society, Series B, 57, 659-672.

Bloomfield, P., 1973, An exponential model in the spectrum of a scalar time series, Biometrika 60, 217-226.

Campbell, J.Y. and R.J. Shiller (1987), "Cointegration and tests of present value models", Journal of Political Economy, 95, 1062-1088.

Campbell, J.Y., and R.J. Shiller, 1998, Valuation Ratios and the Long-Run Stock Market Out-look. The Journal of Portfolio Management, vol. 24, 11-26.

Campbell, J.Y., and R.J. Shiller, 2001, Valuation Ratios and the Long-Run Stock Market Out-look: An Update. Cowles Foundation Discussion Paper No. 1295, Yale University.

Caporale, G.M. and L.A. Gil-Alana (2004), "Fractional cointegration and tests of present value models", Review of Financial Economics, 13, 3, 245-258.

Crowder, W. (1996), "A note on cointegration and international capital market efficiency. A reply", Journal of International Money and Finance 15, 661-664.

Cunado, J., L.A. Gil-Alana and F. Perez de Gracia (2005), A test for ractional bubbles int he Nasdaq stock index. A fractionally integrated approach, Journal of Banking and Finance 29, 10, 2633-2654.

Dahlhaus. R., (1989), "Efficient parameter estimation for self-similar processes", Annals of Statistics 17, 1749-1766.

Demetrescu, M., V. Kuzin and U. Hassler, 2008, Long memory testing in the time domain, Econometric Theory 24, 176-215.

Dolado, J.J. and F. Marmol (1996) Efficient estimation of cointegrating relationships among higher orders and fractionally integrated processes. Banco de España, Servicio de Estudios, Madrid.

Durre, A. and P. Giot, 2007, An international analysis of earnings, stock prices and bond yields, Journal of Business Finance and Accounting 34, 3-4, 613-641.

Engle, R.F. and C.W.J. Granger (1987) Cointegration and error correction model. Representation, estimation and testing. Econometrica 55, 251-276.

Gil-Alana, L.A. (2004) The use of the model of Bloomfield (1973) as an approximation to ARMA processes in the context of fractional integration, Mathematical and Computer Modelling 39, 429-436.

Gil-Alana, L.A. and B .Henry (2003) Fractional integration and the dynamics of UK unemployment, Oxford Bulletin of Economics and Statistics 65, 221-240.

Gil-Alana, L.A. and P.M. Robinson (1997) Testing unit roots and other nonstationary hypotheses in macroeconomic time series, Journal of Econometrics 80, 2, 241-268.

Granger, C.W.J. (1981) Some properties of time series data and their use in econometric model specification. *Journal of Econometrics* 16, 121-130.

Granger, C.W.J. and A. Weiss (1983) Time series analysis of error-correcting models, Studies in Econometrics, Time Series and Multivariate Statistics, New York Academic Press, 255-278.

Hakkio, G. and M. Rush (1989), "Market efficiency and cointegration. An application to the Sterling and Deutschmark exchange markets", Journal of International Money and Finance 8, 75-88.

Han, H.L. (1996), "Cointegration and tests of a present value model in the stock market", *Applied Economics*, 28, 267-272.

Hannan, E.J. (1963) Regression for time series, In M. Rosenblatt (ed.), Time Series Analysis, 17-37, New York, John Wiley.

Hualde, J. and P.M. Robinson (2007) Root-n-consistent estimation of weak fractional cointegration. *Journal of Econometrics* 140, 450-484.

Johansen, S. (1988) Statistical analysis of cointegrating vectors. *Journal of Economics Dynamics and Control* 12, 231-254.

Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551-1580.

Kew, H. and D. Harris, 2009, Heteroskedasticity robust testing for a fractional unit root, Econometric Theory 25, 1734-1753.

Koop, G., 1991Cointegration tests in present value relationships. A Bayesian look at the bivariate properties of stock prices and dividends, Journal of Econometrics 49, 105-139.

Lobato, I. and C. Velasco, 2007, Efficient Wald tests for fractional unit roots, Econometrica 75, 575-589.

Marinucci, D. and P.M. Robinson (2001) Semiparametric fractional cointegration analysis, Journal of Econometrics 105, 225-247.

Narayan, P.K. and R. Smyth (2005), Cointegration of stock markets between New Zealand, Australia and the G-& economies: searching for co-movement under structural change, Australian Economic Papers, 44, 3, 231-247.

Park, J.Y. (1992) Canonical cointegrating regression. *Econometrica* 60, 119-144.

Pereira-Garmendia, D., (2010), Inflation, real stock prices and earnings. Friedman was right, Working Paper, University Pompeu Fabra, Barcelona, Spain.

Philips, T.K., (1999) Why do valuation ratios forecast long-run equity returns?, Journal of Portfolio Management pp. 39–44.

Phillips, P.C.B. (1991a) Optimal inference in cointegrating systems. *Econometrica* 59, 283-306.

Phillips, P.C.B. (1991b) Spectral regresssions for cointegrated time series, in W.A. Barnett, J. Powell and G. Tauchen (eds.) Nonparametric and semiparametric methods in econometrics and statistics, Cambridge, Cambridge University Press.

Phillips, P.C.B. and S.N. Durlauf (1986) Multiple time series regressions with integrated processes. *Review of Economic Studies* 53, 473-495.

Phillips, P.C.B. and B.E. Hansen (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99-125.

Phillips, P.C.B. and M. Loretan (1991) Estimating long run economic equilibria. *Review of Economic Studies* 58, 407-436.

Phillips, P.C.B. and Shimotsu, K., 2004. Local Whittle estimation in nonstationary and unit root cases. Annals of Statistics 32, 656-692.

Rangvid, J. (2001), "Increasing convergence among European stock markets? A recursive common stochastic trends analysis, Economics Letters 71, 383-389

Richards, A.J. (1995), "Comovements in national stock market returns. Evidence of predictability but not cointegration, Journal of Monetary Economics 36, 631-654.

Robinson, P.M. (1994a), "Efficient tests of nonstationary hypotheses", *Journal of the American Statistical Association*, 89, 1420-1437.

Robinson, P.M. (1994b) Semiparametric analysis of long memory time series. *Annals of Statistics* 22, 515-539.

Robinson, P.M., 1995a, Gaussian semi-parametric estimation of long range dependence, Annals of Statistics 23, 1630-1661.

Robinson, P.M., 1995b, Log-periodogram regression of time series with long range dependence, Annals of Statistics 23, 1048-1072.

Robinson, P.M. and M. Henry, 1999, Long and short memory conditional heteroskedasticity in estimating the memory in levels, Econometric Theory 15, 299-336.

Robinson, P.M. and J. Hualde (2003) Cointegration in fractional systems with unknown integration orders. *Econometrica* 71, 1727-1766.

Robinson, P.M. and Y. Yajima (2002), Determination of cointegrating rank in fractional systems, Journal of Econometrics 106, 217-241.

Saikkonen, P. (1991) Asymptotically efficient estimation of cointegrating regressions. *Econometric Theory* 7, 1-21.

Shimotsu, K. and P.C.B. Phillips, 2005. Exact local Whittle estimation of fractional integration. Annals of Statistics 33, 1890-1933.

Sowell, F., 1992, Maximum likelihood estimation of stationary univariate fractionally integrated time series models, Journal of Econometrics 53, 165.188.

Stock, J.H. and M.W. Watson (1993) A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783-820.

Subramanian, U. (2008), Cointegration of stock markets of East Asia, European Journal of Economics, Finance and Administrative Sciences, 14, 84-92.

Velasco, C., 1999b. Gaussian semiparametric estimation of nonstationary time series. Journal of Time Series Analysis 20, 87-127.

Velasco, C. and P.M. Robinson, 2000, Whitle pseudo maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association 95, 1229-1243.

Tables and Figures

Table 1: Estimates of d in a model with white noise disturbances

	No Regressors	An intercept	A linear time trend
STOCK PRICES	1.169 (1.134, 1.208)	1.169 (1.135, 1.209)	1.170 (1.135, 1.209)
DIVIDENDS	1.906 (1.874, 1.941)	1.951 (1.916, 1.988)	1.951 (1.916, 1.988)
EARNINGS	1.855 (1.806, 1.910)	1.856 (1.806, 1.911)	1.856 (1.806, 1.911)
CONSUMER PRICE INDEX	1.210 (1.185, 1.241)	1.396 (1.350, 1.454)	1.401 (1.354, 1.456)
LONG INTEREST RATE	1.111 (1.070, 1.157)	1.111 (1.070, 1.157)	1.110 (1.070, 1.156)
REAL STOCK PRICES	1.156 (1.121, 1.195)	1.161 (1.126, 1.201)	1.161 (1.126, 1.201)
REAL DIVIDENDS	1.311 (1.279, 1.346)	1.505 (1.470, 1.544)	1.505 (1.470, 1.544)
REAL EARNINGS	1.756 (1.713, 1.803)	1.825 (1.779, 1.877)	1.825 (1.779, 1.877)
PRICE /EARNING RATIO	1.237 (1.198, 1.282)	1.494 (1.449, 1.542)	1.494 (1.449, 1.542)

The values in parentheses refer to the 95% confidence band of the non-rejection values of d using Robinson's (1994a) parametric tests.

Table 2: Estimates of d in a model with Bloomfield-type disturbances

	No Regressors	An intercept	A linear time trend
STOCK PRICES	1.052 (1.000, 1.102)	1.052 (1.001, 1.102)	1.052 (1.001, 1.103)
DIVIDENDS	1.033 (0.987, 1.083)	1.037 (0.986, 1.087)	1.037 (0.985, 1.088)
EARNINGS	1.568 (1.499, 1.649)	1.569 (1.492, 1.653)	1.569 (1.492, 1.653)
CONSUMER PRICE INDEX	1.175 (1.148, 1.206)	1.187 (1.160, 1.211)	1.195 (1.173, 1.224)
LONG INTEREST RATE	0.909 (0.864, 0.964)	0.908 (0.863, 0.964)	0.909 (0.864, 0.964)
REAL STOCK PRICES	1.033 (0.981, 1.083)	1.037 (0.991, 1.087)	1.037 (0.991, 1.088)
REAL DIVIDENDS	1.339 (1.272, 1.419)	1.448 (1.388, 1.521)	1.448 (1.388, 1.521)
REAL EARNINGS	1.599 (1.510, 1.671)	1.600 (1.517, 1.681)	1.600 (1.517, 1.681)
PRICE /EARNING RATIO	1.135 (1.060, 1.231)	1.269 (1.171, 1.398)	1.269 (1.171, 1.398)

The values in parentheses refer to the 95% confidence band of the non-rejection values of d using Robinson's (1994a) parametric tests.

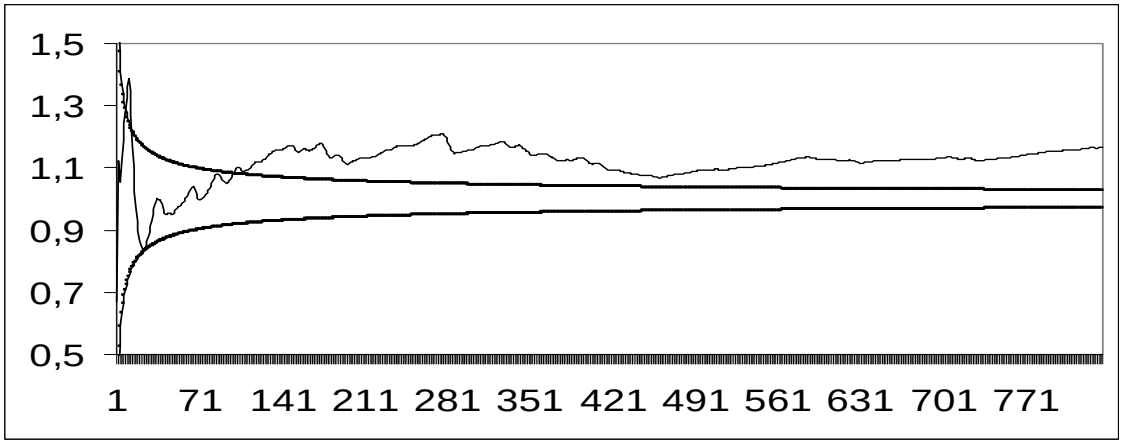
Table 3: Estimates of d in a model with seasonal AR(1) disturbances

	No Regressors	An intercept	A linear time trend
STOCK PRICES	1.169 (1.135, 1.208)	1.170 (1.135, 1.210)	1.170 (1.135, 1.210)
DIVIDENDS	1.902 (1.873, 1.935)	1.953 (1.921, 1.988)	1.953 (1.921, 1.988)
EARNINGS	1.875 (1.830, 1.926)	1.878 (1.832, 1.928)	1.878 (1.833, 1.928)
CONSUMER PRICE INDEX	1.188 (1.161, 1.220)	1.374 (1.326, 1.431)	1.378 (1.332, 1.434)
LONG INTEREST RATE	1.111 (1.071, 1.157)	1.111 (1.071, 1.157)	1.110 (1.071, 1.157)
REAL STOCK PRICES	1.155 (1.119, 1.194)	1.161 (1.125, 1.201)	1.161 (1.125, 1.201)
REAL DIVIDENDS	1.311 (1.279, 1.346)	1.505 (1.469, 1.544)	1.505 (1.469, 1.544)
REAL EARNINGS	1.742 (1.704, 1.787)	1.836 (1.793, 1.883)	1.838 (1.795, 1.885)
PRICE /EARNING RATIO	1.234 (1.194, 1.278)	1.491 (1.447, 1.539)	1.491 (1.447, 1.540)

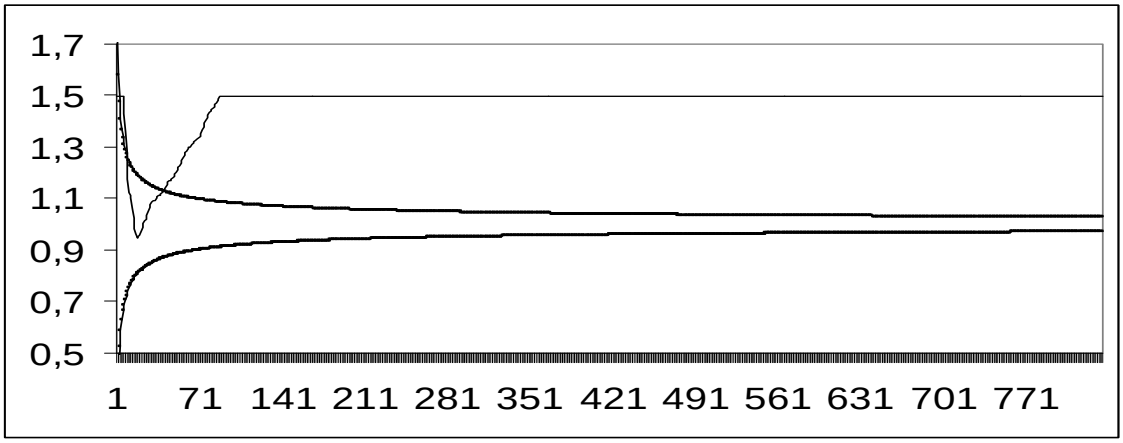
The values in parentheses refer to the 95% confidence band of the non-rejection values of d using Robinson's (1994a) parametric tests.

Figure 1: Estimates of d based on the semiparametric estimate of Robison (1995)

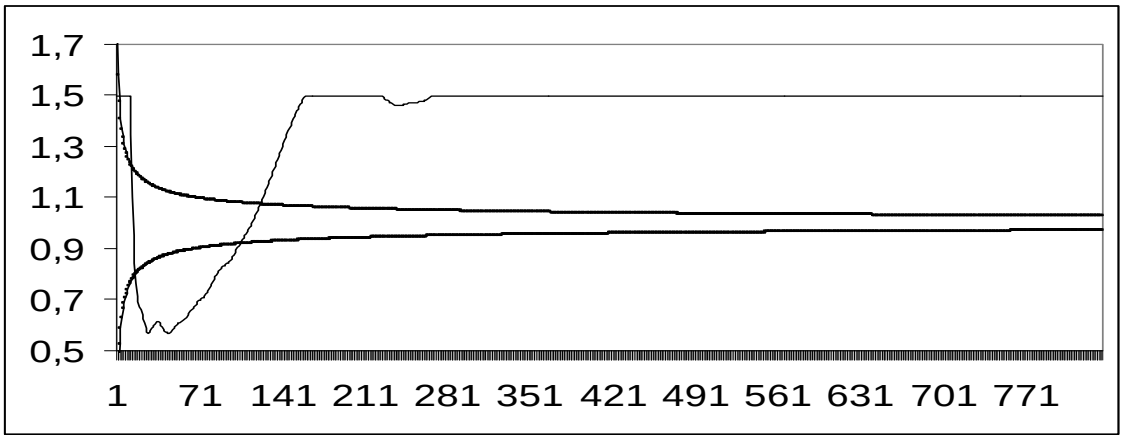
Stock prices



Dividends



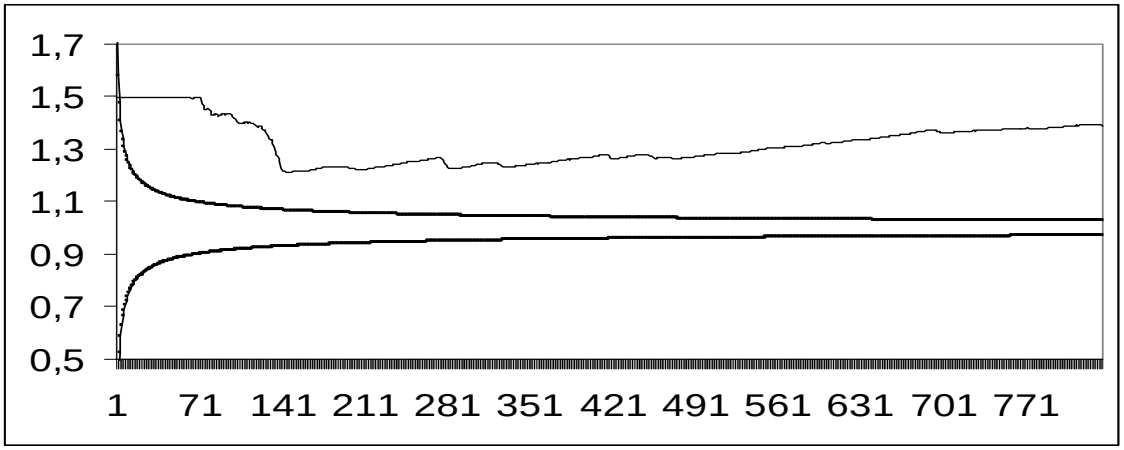
Earnings



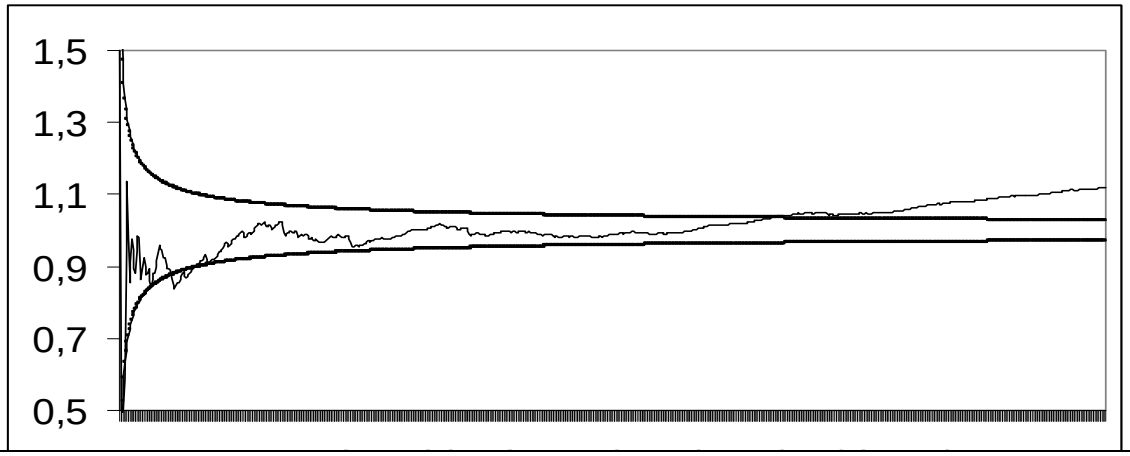
(cont.)

Figure 1: Estimates of d: semiparametric estimate of Robison (1995)-cont.

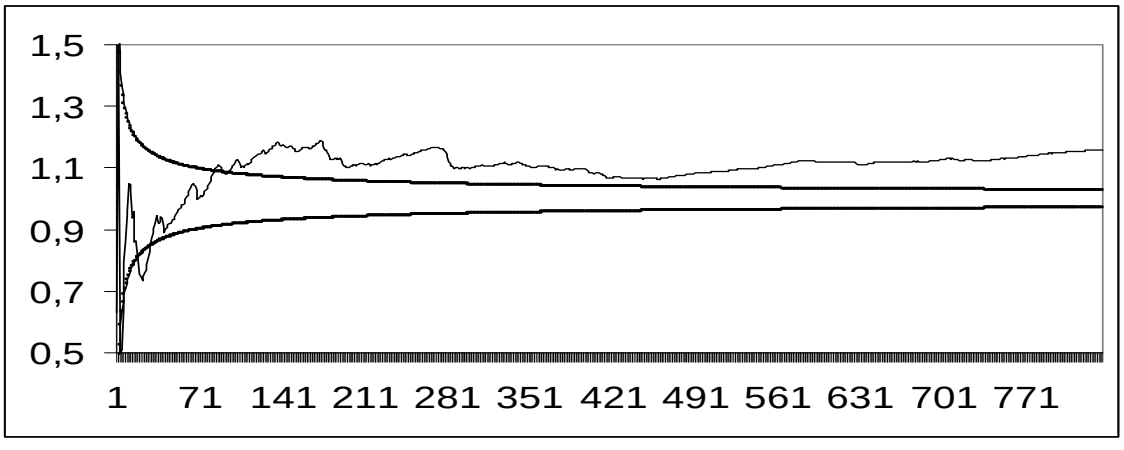
Consumer Price Index



Long interest rates

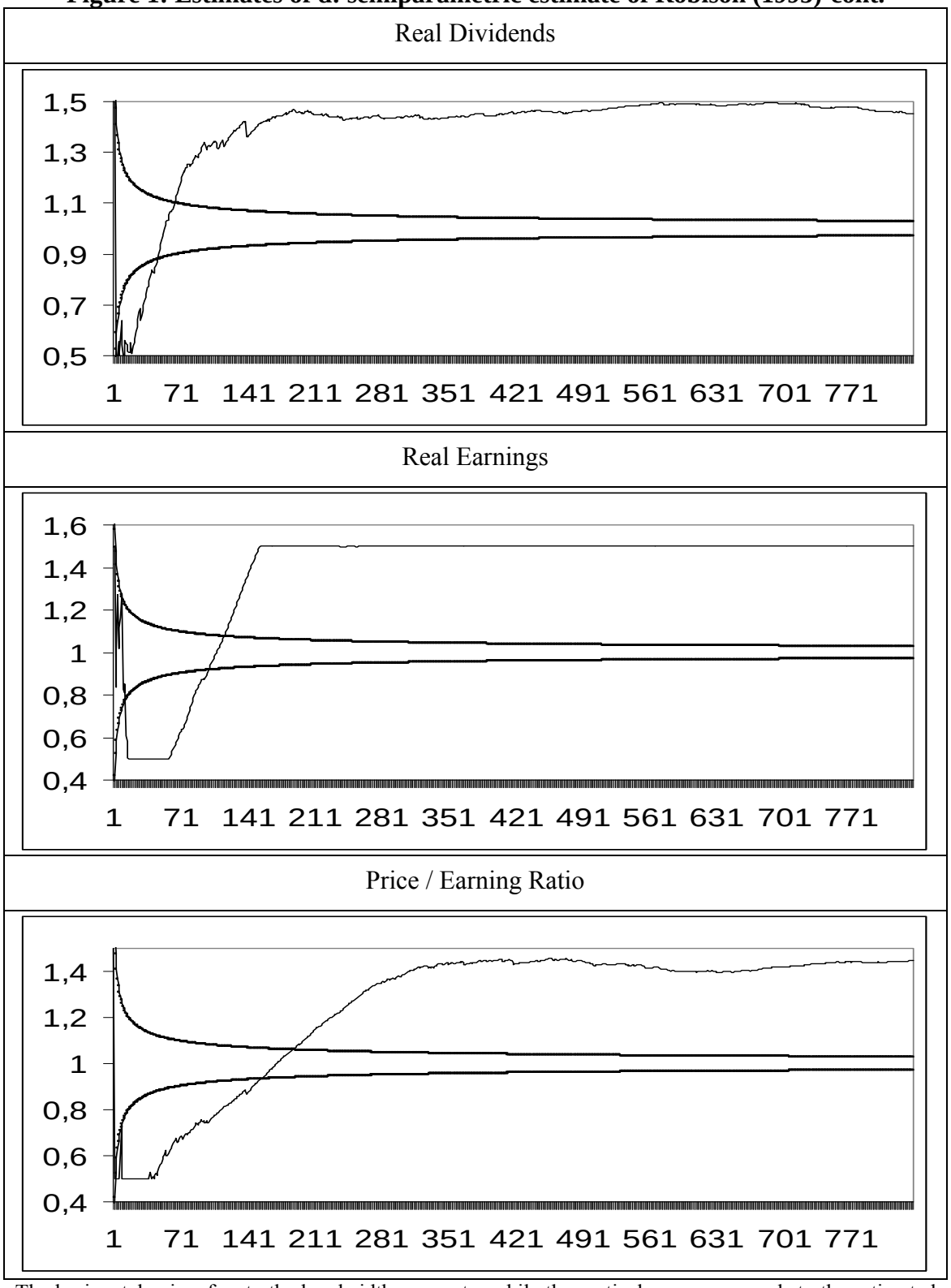


Real stock prices



(cont.)

Figure 1: Estimates of d: semiparametric estimate of Robison (1995)-cont.



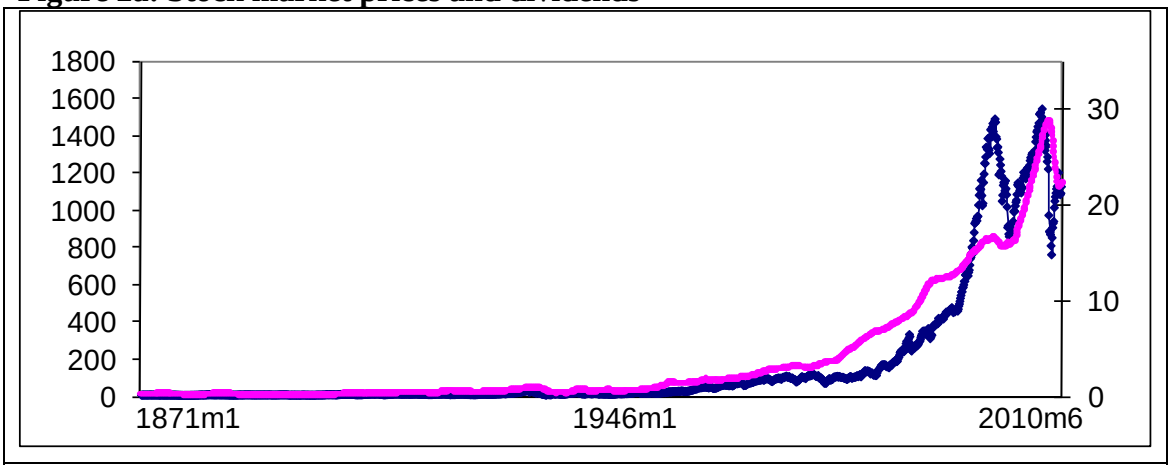
The horizontal axis refers to the bandwidth parameter while the vertical one corresponds to the estimated values of d. We report the estimates of d along with the 95% confidence band of the I(1) hypothesis.

Table 4: Estimates of d using Robinson's (1995) semiparametric method for different bandwidth numbers

	25	41	100	200	300	500
STOCK PRICES	0.850*	0.953*	1.004*	1.121	1.158	1.092
DIVIDENDS	1.021*	1.105*	1.500	1.500	1.500	1.500
EARNINGS	0.589	0.580	0.875	1.500	1.500	1.500
CONSUMER PRICE INDEX	1.500	1.500	1.417	1.228	1.235	1.278
LONG INTEREST RATE	0.893*	0.895*	0.983*	0.958*	0.990*	1.013*
REAL STOCK PRICES	0.768	0.888*	1.071*	1.107	1.099	1.086
REAL DIVIDENDS	0.538	0.896*	1.326	1.455	1.438	1.464
REAL EARNINGS	0.500	0.500	0.933*	1.500	1.500	1.500
PRICE /EARNING RATIO	0.500	0.500	0.745	1.041*	1.377	1.431
95% Confidence Interval	(0.835, 1.164)	(0.871, 1.128)	(0.917, 1.082)	(0.941, 1.058)	(0.952, 1.047)	(0.963, 1.036)

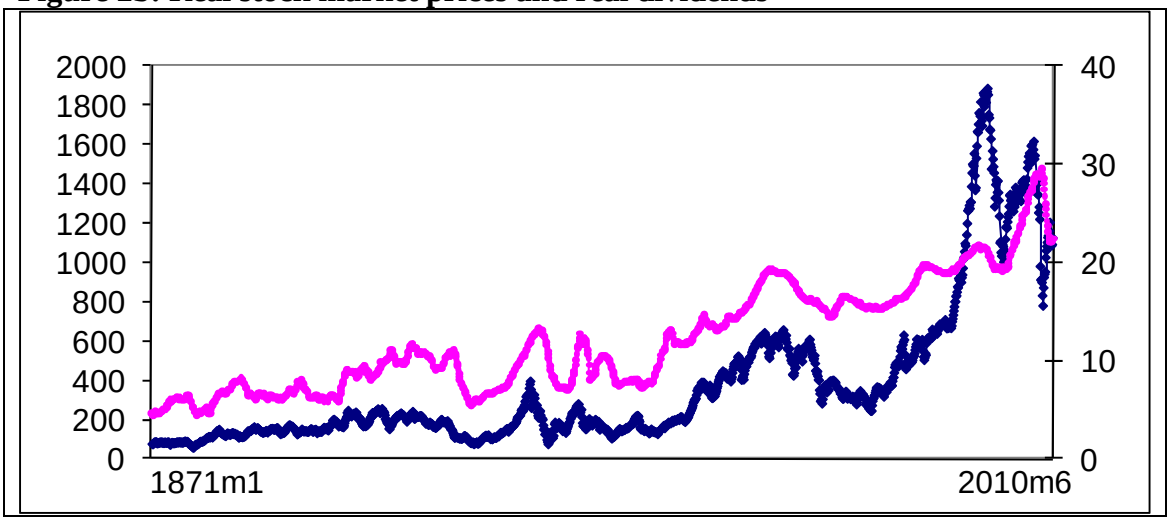
* indicates that the I(1) hypothesis cannot be rejected at the 5% level.

Figure 2a: Stock market prices and dividends



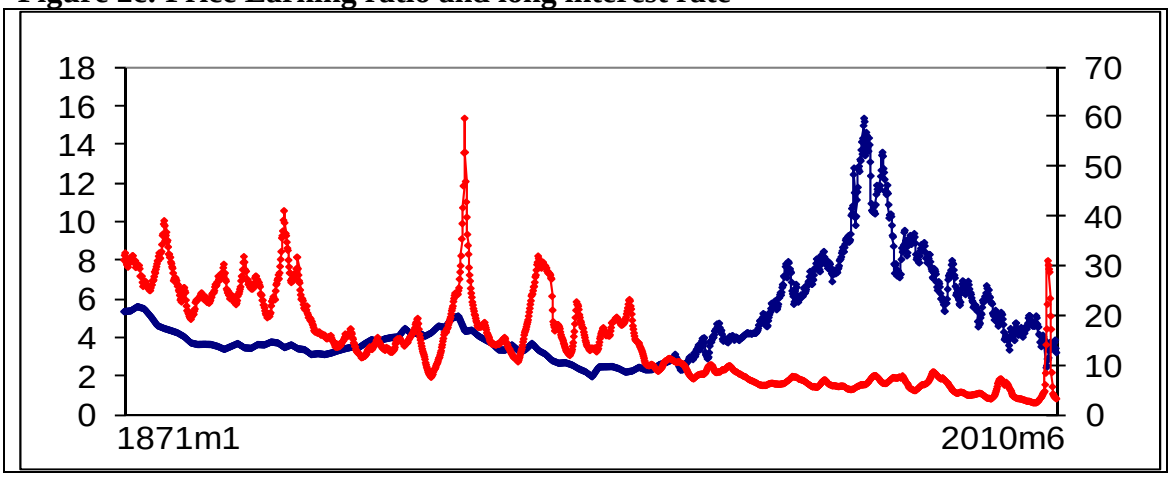
The thick line refers to the stock market prices and the thin one is for dividends.

Figure 2b: Real stock market prices and real dividends



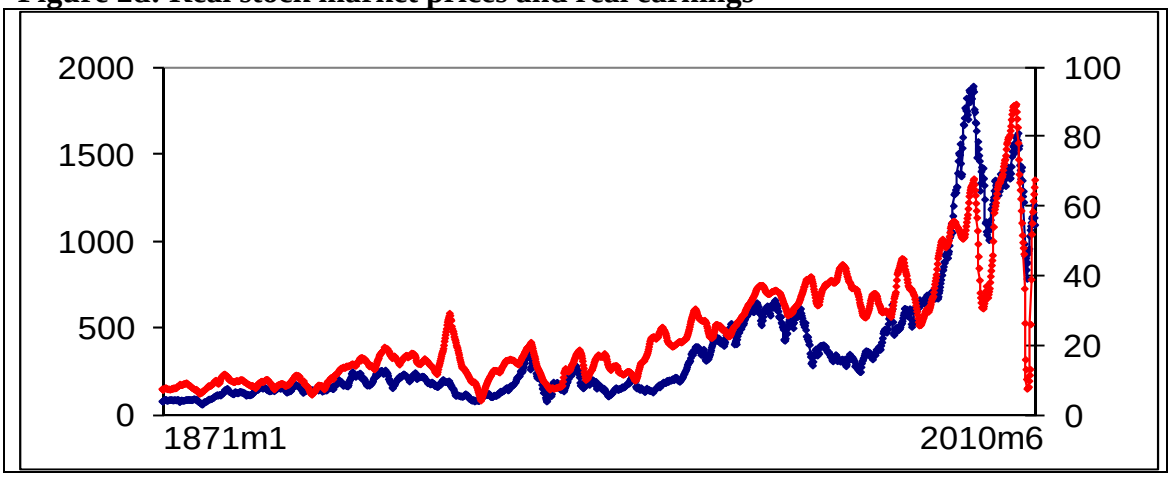
The thick line refers to real stock market prices and the thin one to real dividends.

Figure 2c: Price Earning ratio and long interest rate



The thick line refers to the long-term interest rate and the thin one is for the price earning ratio.

Figure 2d: Real stock market prices and real earnings



The thick line refers to the real stock market prices and the thin one is for real earnings.

Table 5: Estimates of d using Robinson's (1995b) log-periodogram semiparametric method for different values of l and fixed m = $(T)^{0.5}$

$m = T^{0.5}$	l = 0	l = 1	l = 2	l = 3	l = 4	l = 5
STOCK PRICES	1.080	1.070	1.061	1.041	1.018	1.017
DIVIDENDS	1.095	1.048	1.026	1.037	1.134	1.222
LONG INTEREST RATE	0.956	1.010	0.972	0.914	0.832	0.803
REAL STOCK PRICES	0.972	1.000	1.085	1.077	1.018	1.016
REAL DIVIDENDS	0.822	0.851	0.981	0.975	0.996	0.997
REAL EARNINGS	0.970	1.009	1.073	1.099	1.129	1.162
REAL EARNINGS	0.279	0.128	0.059	-0.078	-0.142	-0.082
PRICE /EARNING RATIO (*)	0.913	0.931	0.945	0.920	1.010	1.127
PRICE /EARNING RATIO (**)	0.484	0.606	0.576	0.589	0.637	0.645

(*) and (**) indicates that the results are based on the original and first differenced data respectively.

Table 6: Testing the homogeneity in the order of integration (Robinson and Yajima, 2002)

$m = T^{0.5}$	l = 0	l = 1	l = 2	l = 3	l = 4	l = 5
Stock prices / Dividends	-0.145	0.214	0.340	0.038	-1.125	-1.970
Real stock prices / Real dividens	1.471	1.455	1.009	0.990	0.215	0.185
P.E.R. / Long interest rates	-0.425	-0.776	-0.813	-1.617	-1.527	-0.675
Real stock prices / Real earnings	0.580	0.676	1.356	1.520	0.078	-1.077

In all cases we employ h(n) chosen as b^{-5-2i}, i=1,2,3,4 and 5.

Table 7: Hausman test for no cointegration (Marinucci and Robinson, 2001)

$\mathbf{m = T^{0.5}}$	H	s = 25	s = 41	s = 50
Stock prices / Dividends	H_{as}	4.205*	0.239	0.102
	H_{bs}	2.420	1.260	1.040
	$\hat{d}_*$	0.916	1.088	1.077
Real stock prices / Real dividends	H_{as}	26.499*	11.469*	8.880*
	H_{bs}	13.520*	2.259	0.810
	$\hat{d}_*$	0.721	0.898	0.936
P.E.R. / Long interest rates	H_{as}	24.780*	16.457*	13.209*
	H_{bs}	28.728*	20.664*	32.262*
	$\hat{d}_*$	0.593	0.721	0.688
Real stock prices / Real earnings	H_{as}	71.520*	68.502*	74.649*
	H_{bs}	68.679*	64.952*	70.560*
	$\hat{d}_*$	0.487	0.628	0.653

$X_1^2(5\%) = 3.84$. * indicates rejection of the null hypothesis of no cointegration at the 5% level.

Table 8: Coefficient estimates in a fractional cointegration setting using different methods

	Stock prices / Dividends	Real stock prices / Real dividends	P.E.R. / Long interest rates	Real stock prices / Real earnings
ν (OLS) time domain	50.074	35.514	2.173	17.636
ν (OLS) freq. domain	69.214	46.539	1.707	32.267
ν (NBLS) (m = 25)	40.920	31.347	1.507	7.532
ν (NBLS) (m = 41)	37.650	29.506	1.149	10.494
ν (NBLS) (m = 100)	38.802	29.929	0.857	20.528
ν (NBLS) (m = 200)	43.209	32.361	1.241	20.577
ν (NBLS) (m = 300)	47.291	34.668	1.176	22.777
ν (NBLS) (m = 400)	51.397	36.957	1.263	24.523
ν (NBLS) (m = 500)	55.518	39.188	1.382	26.319

Table 9: Estimates of d on the residuals using the log-periodogram estimate of Robinson (1995b)

$m = T^{0.5}$ $l = 0$	Stock prices / Dividends	Real stock prices / Real dividends	P.E.R. / Long interest rates	Real stock prices / Real earnings
ν (OLS) time domain	0.775	1.082	0.811	0.723
ν (OLS) freq. domain	0.633	0.976	0.828	0.526
ν (NBLS) (m = 25)	1.012	1.112	0.837	1.129
ν (NBLS) (m = 41)	1.126	1.132	0.860	1.042
ν (NBLS) (m = 100)	1.083	1.128	0.878	0.600
ν (NBLS) (m = 200)	0.942	1.103	0.855	0.598
ν (NBLS) (m = 300)	0.832	1.087	0.859	0.560
ν (NBLS) (m = 400)	0.754	1.077	0.853	0.551
ν (NBLS) (m = 500)	0.708	1.070	0.846	0.532

Table 10: Estimates of d on the residuals using the log-periodogram estimate of Robinson (1995b)

$m = T^{0.5}$ $l = 2$	Stock prices / Dividends	Real stock prices / Real dividends	P.E.R. / Long interest rates	Real stock prices / Real earnings
ν (OLS) time domain	0.991	1.273	0.650	0.805
ν (OLS) freq. domain	0.617	1.127	0.697	0.430
ν (NBLS) (m = 25)	1.260	1.304	0.725	1.283
ν (NBLS) (m = 41)	1.386	1.326	0.781	1.190
ν (NBLS) (m = 100)	1.337	1.322	0.827	0.623
ν (NBLS) (m = 200)	1.186	1.294	0.766	0.620
ν (NBLS) (m = 300)	1.066	1.277	0.777	0.557
ν (NBLS) (m = 400)	0.956	1.269	0.763	0.535
ν (NBLS) (m = 500)	0.853	1.263	0.744	0.486

Table 11: Estimates of d on the residuals using the Whittle estimate of Robinson (1995a)

m = 25	Stock prices / Dividends	Real stock prices / Real dividends	P.E.R. / Long interest rates	Real stock prices / Real earnings
ν (OLS) time domain	0.606	0.733	0.500	0.610
ν (OLS) freq. domain	0.608	0.709	0.500	0.500
ν (NBLS) (m = 25)	0.663	0.741	0.500	0.878
ν (NBLS) (m = 41)	0.683	0.744	0.500	0.843
ν (NBLS) (m = 100)	0.676	0.743	0.500	0.515
ν (NBLS) (m = 200)	0.648	0.739	0.500	0.514
ν (NBLS) (m = 300)	0.622	0.735	0.500	0.500
ν (NBLS) (m = 400)	0.599	0.730	0.500	0.500
ν (NBLS) (m = 500)	0.582	0.726	0.500	0.500

Table 12: Estimates of d on the residuals using the Whittle estimate of Robinson (1995a)

$m = T^{0.5} = 41$	Stock prices / Dividends	Real stock prices / Real dividends	P.E.R. / Long interest rates	Real stock prices / Real earnings
ν (OLS) time domain	0.778	0.803	0.612	0.715
ν (OLS) freq. domain	0.781	0.777	0.581	0.500
ν (NBLS) (m = 25)	0.827	0.814	0.567	1.047
ν (NBLS) (m = 41)	0.815	0.819	0.543	0.987
ν (NBLS) (m = 100)	0.811	0.817	0.526	0.623
ν (NBLS) (m = 200)	0.796	0.811	0.549	0.621
ν (NBLS) (m = 300)	0.784	0.805	0.545	0.562
ν (NBLS) (m = 400)	0.775	0.800	0.551	0.521
ν (NBLS) (m = 500)	0.770	0.794	0.559	0.500

Table 13a: Stock prices and dividends

Model	ν	d	γ
u_{1t} is white noise u_{2T} is white noise	52.754	1.161	0.773
u_{1t} is AR(1) u_{2T} is white noise	48.829	1.166	0.795
u_{1t} is white noise u_{2T} is AR(1)	48.829	1.188	0.609
u_{1t} is AR(1) u_{2T} is AR(1)	48.792	1.151	0.631

Table 13b: Real stock prices and real dividends

Model	ν	d	γ
u_{1t} is white noise u_{2T} is white noise	57.435	1.047	0.780
u_{1t} is AR(1) u_{1T} is white noise	52.251	1.036	0.763
u_{1t} is white noise u_{2T} is AR(1)	52.249	0.996	0.526
u_{1t} is AR(1) u_{2T} is AR(1)	52.208	1.159	0.878

Table 13c: P.E.R. and long interest rates

Model	ν	d	γ
u_{1t} is white noise u_{2T} is white noise	-1.566	1.165	0.779
u_{1t} is AR(1) u_{2T} is white noise	0.874	1.053	0.763
u_{1t} is white noise u_{2T} is AR(1)	0.876	1.115	0.527
u_{1t} is AR(1) u_{2T} is AR(1)	0.874	1.153	0.877

Table 13d: Real stock prices and real earnings

Model	ν	d	γ
u_{1t} is white noise u_{2T} is white noise	20.253	1.081	0.780
u_{1t} is AR(1) u_{2T} is white noise	13.376	1.150	0.764
u_{1t} is white noise u_{2T} is AR(1)	13.387	0.984	0.521
u_{1t} is AR(1) u_{2T} is AR(1)	13.390	1.152	0.890