Review Article
Spurious Regression

D. Ventosa-Santaulària

Departamento de Economia y Finanzas, Universidad de Guanajuato, DCEA-Campus Marfil Fracc. I, 36250 El Establo, Guanajuato, Gto, Mexico

Correspondence should be addressed to D. Ventosa-Santaulària, daniel@ventosa-santaularia.com

The spurious regression phenomenon in least squares occurs for a wide range of data generating processes, such as driftless unit roots, unit roots with drift, long memory, trend and broken-trend stationarity. Indeed, spurious regressions have played a fundamental role in the building of modern time series econometrics and have revolutionized many of the procedures used in applied macroeconomics. Spin-offs from this research range from unit-root tests to cointegration and error-correction models. This paper provides an overview of results about spurious regression, pulled from disperse sources, and explains their implications.

1. Introduction

During the last 30 years econometric theory has undergone a revolution. In the late seventies, economists and econometricians recognized that insufficient attention was being paid to trending mechanisms and that, in fact, most macroeconomic variables were probably nonstationary. Such an appraisal gave rise to an extraordinary development that substantially modified the way empirical studies in time-series econometrics are carried out. Research in nonstationarity has advanced significantly since the early important papers, such as Granger and Newbold [1], Davidson et al. [2], Hendry and Mizon [3], Plosser and Schwert [4], Bhattacharya et al. [5], and Phillips [6]. Nelson and Plosser [7] asserted that many relevant U.S. macroeconomic time series were governed by a unit root (a random trending mechanism), based on Dickey and Fuller's [8] Unit-root test. Several years later, Perron [9] argued that the trending mechanism in macro variables was deterministic in nature (with some transcendent structural breaks). The debate continues between "unit rooters" and "deterministic trenders", though there is very general consensus as to the presence of a trending mechanism in the levels of most macroeconomic series. In the words of Durlauf and Phillips [10]: "Traditional Analyses of Economic time series frequently rely on the assumption that the time series in question are stationary, ergodic processes [...]. However,

the assumptions of the traditional theory do not provide much solace to the empirical worker. Even casual examination of such time series as GNP reveals that the series do not possess constant means."

Econometrics should work hand-in-hand with economic theory by providing it with the tools it requires to understand economic activity. The modeling of such mechanisms is thus a major goal of time series econometrics (that said, it should be acknowledged that neither unit roots nor deterministic trends are able to model satisfactorily most series; they are what Phillips [11] labeled "heroically naive" concepts). Spurious regression can be considered as having played a fundamental role in this development. To understand it, we paraphrase Granger et al. [12], who provide an illuminating definition (Phillips [6] showed analytically that, when regressing two independent stochastic trends, the estimates of the regression coefficient do not converge to their real value of zero (particularly noteworthy among the various spin-offs of the aforementioned studies are the Error Correction Model (first proposed by Sargan [13] as a link between static equilibrium economic theory and dynamic empirical models, and further developed by Hendry and Anderson [14], Davidson et al. [2], Hendry and Mizon [3], and Hendry et al. [15]) and the entire theory of cointegration (first proposed by Granger [16], Granger and Weiss [17], and Engle and Granger [18]. The linking of Cointegration with Error Correction Models can also be found in these articles). Spurious regression has contributed to the general improvement in the level of empirical work). Phillips's [6] results are detailed in a simple case in Appendix A): "A spurious regression occurs when a pair of independent series, but with strong temporal properties, is found apparently to be related according to standard inference in a Least Squares regression."

Phillips [19] presented a counterargument on the usefulness of spurious regression: trend specifications are just coordinate systems for representing behavior over time. Phillips argues that even if the series are statistically independent, when they include a trending mechanism in their DGP, they admit a regression representation (even in the absence of cointegration). This is in sharp contrast to the usual concept of spurious regression. We usually conceive this phenomenon as the statistical identification of a commonality of trending mechanisms. Phillips ventures that such results—the "spurious"—constitute an adequate representation of the data. His main result applies for regressions among stochastically trended series on time polynomials as well as regressions among independent random walks. Phillips [19] proves that Brownian motions can be represented by deterministic functions of time with random coefficients. Given that standardized discrete time series with a unit root (hereinafter UR) converge weakly to Brownian motion processes, it is argued that deterministic time functions may be used to model them. Such representations include polynomial trends, trend breaks as well as sinusoidal trends; it is also proved that a stochastic process can represent an arbitrary deterministic function on a particular interval, so a regression of a UR process on an independent UR process is thus also a valid representation of the data. In both cases, the t-statistics diverge at rate $T^{1/2}$, which is consistent since such parameterization reflects a partial—though correct—specification of the DGP. One of the most significant conclusions of Phillips concerns the long-standing debate of UR versus Trend Stationarity. To quote Phillips: "[...] Our results show that such specifications (Trend stationary processes) are not, in fact, really alternatives to a UR model at all. Since the UR processes have limiting representations entirely in terms of these functions (deterministic), it is apparent that we can mistakenly "reject" a UR model in favor of a trend "alternative" when in fact that alternative model is nothing other than an alternate representation of the UR process itself ."

This perspective has the virtue of allowing variables with different trending mechanism (deterministic or stochastic) to be related without being limited to the somewhat restrictive case of cointegration. Phillips advances this as an appropriate approach to study stochastically unbalanced relationships such as the ones that may arise between variables such as interest rates, inflation, money stock and GDP (for further detail, see Phillips [20]).

To the best of our knowledge, little has been done in the way of bringing the most important works in this field together, treating them in any kind of standardized way, making connections between them, and of any real study of the profound implications for economics they might indicate. This article aims to rectify this situation.

2. Appraisal of the Spurious Phenomenon

Much progress has been made with Least Squares statistical inference since it was first proposed more than two centuries ago as a means of estimating the course of comets [21]. Theoretical developments in econometrics address the nonexperimental nature of economic data sets. Least Squares (LS) offers a trade-off between simplicity and powerful inference. Nevertheless, LS has certain limitations, such as potential confusion between correlation and causality, and used unwisely may produce misleading evidence. Statisticians and econometricians had been aware of the "spurious phenomenon" since Yule [22], and Pearson [23], (for excellent reviews of these works, see Hendry and Morgan [24], and Aldrich [25]). These results led to the common expertise in the time-series field that indicated the need to differentiate potentially nonstationary series when using these to run regressions or detrending these by fitting trend lines estimated with LS. See Morgan [26].

There are many examples of spurious regression. Some of these are commented on in Phillips [19], where we discover the implausible relationship between "the number of ordained ministers and the rate of alcoholism in Great Britain in the nineteenth century"; the equally "remarkable relationship" presented in Yule [27] concerning the "proportion of Church of England marriages to all marriages and the mortality rate over the period 1866–1911"; the "strange relationship" between price level and cumulative rainfall in the UK, which was advanced as a curious alternative version of quantitative theory by Hendry [28]. Plosser and Schwert [4] presented another example of nonsense correlation when they proposed their quantity theory of sunspots. The main argument is that the log of nominal income can be explained by way of the log of accumulative sunspots. Not only did they find statistically significant estimates, but the goodness of fit, measured with the R^2, is quite high: 0.82 (a variant of this example, used by the authors to demonstrate the danger of nonsense correlation was taken seriously 100 years ago, by Jevons [29]). Granger and Newbold [1] computed a Monte Carlo Experiment where a number of regressions, specified as (1), were run using simulated variables, each perfectly independent of the others:

$$y_t = \alpha + \beta x_t + u_t, \tag{2.1}$$

where $t = 1, \ldots, T$, being T the sample size. The variables x_t and y_t are independent. Under standard regularity conditions, LS delivers no evidence of a linear relationship between y and x. In particular, $\hat{\beta}$ should be statistically equal to zero. Nevertheless, Granger and Newbold's [1], they were generated as $I(1)$ processes, ($I(d)$ notation, for d an integer, refers to the number of differences to be performed so the variable becomes stationary). usually referred to as random walks, that is, UR processes, but found estimated parameters statistically different

from zero, with its associated t-ratio $t_{\hat{\beta}} = \hat{\beta}/\hat{\sigma}_{\hat{\beta}}$, unusually high (we could technically refer to this as Type II spurious regression. This distinction was first proposed by Chiarella and Gao (2002) as an analogy of the classical type I and type II errors in statistics: (1) Type I spurious regression: rejection of a true relationship, and, (2) Type II spurious regression: acceptance of a false relationship. Most of the literature refers exclusively to Type II spurious regression, which we will simply refer to as spurious regression). Phillips [6] provided a theoretical framework that explained the causes of the phenomenon of spurious regression. In short, it is fair to say that standard LS inference can only be drawn when the variables are stationary. Even with stationary but highly persistent variables, spurious regression can occur when the standard errors used in the t-ratio are inconsistent. Ferson et al. [30] provide examples in financial economics. One extremely important exception to this is the case of cointegration. Even if the series are stochastically trending, when the trend is common to both series the LS regression then works particularly well in the sense that estimates converge in probability to their true value at a rate faster than T, but have a nonstandard distribution [31].

3. Data Generating Processes

Research on spurious regression has been making use of increasingly complex Data Generating Processes (DGPs). Table 1 provides a summary of those appearing in this survey: where u_{wt} are independent innovations obeying in Phillips [6, Assumption 1], e_{wt} is an i.i.d white noise with mean zero and variance σ_e^2, and DU_{iwt}, DT_{iwt} are dummy variables allowing changes in the trend's level and slope respectively, that is, $DU_{iwt} = \mathbf{1}(t > T_{b_{iw}})$ and $DT_{iwt} = (t - T_{b_{iw}})\mathbf{1}(t > T_{b_{iw}})$, where $\mathbf{1}(\cdot)$ is the indicator function, and $T_{b_{iw}}$ is the unknown date of the ith break in w. We denote the break fraction as $\lambda_{iw} = (T_{b_{iw}}/T) \in (0,1)$, where T is the sample size; $d \in (-1/2, 3/2)$. Only DGPs 1, 2 and 10 (for $d < 0.5$) satisfy the weak stationarity definition. The remaining DGPs generate nonstationary series (notice that DGP 6 is a special case of DGP 10; nevertheless, both will be treated as if they were different, given that each has been considered independently in the literature). FI processes deserve further discussion; contrary to regular $ARMA(p,q)$ processes—made popular by Box and Jenkins in the 1970s—such as DGP 1, whose autocorrelation function decays at an exponential rate (short memory), $FI(d)$ processes have an autocorrelation function that decays at a hyperbolic rate (long memory).

There are a number of empirical examples in time series in which dependence falls slowly across time. This phenomenon, known as Long Memory or Long-Range dependence, was observed in geophysical data, such as river flow data [32] and in climatological series [33], as well as in economic time series [34]. In three important papers [35–37], the authors extended these processes to provide more flexible low-frequency or Long Memory behavior by considering $I(d)$ processes with noninteger values of d. As pointed out by Granger and Joyeux [36], *"It was standard procedure to consider differencing time series to achieve stationarity"*—thus obtaining a form of the series that can be identified as an $ARMA$ model—however, *"Some econometricians were reluctant to this technique, believing that they may be losing information, by zapping out the low frequency components."* But using infinite variance series without differencing them was also a source of difficulties at that time. Fractional integration encompasses $ARMA$ models for $d = 0$, and $ARIMA$ models for $d = 1$. The process is stationary and ergodic when $d \in (-1/2, 1/2)$; nonstationary but mean-reverting (that is, it returns to its equilibrium or long-run behavior after any random shock). when $1/2 < d < 1$, and nonstationary and mean averting when $d \geq 1$.

Table 1: The DGPs for $w_t = x_t, y_t, z_t$. *Note*: TS, br and dr stand for trend-stationarity, breaks, and drift, respectively.

#	Name	Model
1	$MA(q_w)$ or	$w_t = \sum_{i=1}^{q_w} \Theta_{iw}\epsilon_{wt-i}$ or
	$AR(p_w)$	$w_t = \sum_{i=1}^{p_w} \phi_{wi}w_{t-i} + \epsilon_{wt}$ (stationary)
2	$I(0)$	$w_t = \mu_w + u_{wt}$
3	$I(0) + br$	$w_t = \mu_w + \sum_{i=1}^{N_w} \theta_{iw}DU_{iwt} + u_{wt}$
4	TS	$w_t = \mu_w + \beta_w t + u_{wt}$
5	$TS + br$	$w_t = \mu_w + \sum_{i=1}^{N_w} \theta_{iw}DU_{iwt} + \beta_w t + \sum_{i=1}^{M_w} \gamma_{iw}DT_{iwt} + u_{wt}$
6	$I(1)$	$\Delta w_t = u_{wt}$
7	$I(1) + dr$	$\Delta w_t = \mu_w + u_{wt}$
8	$I(1) + dr + br$	$\Delta w_t = \mu_w + \sum_{i=1}^{N_w} \theta_{iw}DU_{iwt} + u_{wt}$
9	$I(k)$	$\Delta^k w_t = u_{wt}$ for $k = 2, 3, \ldots$
10	$FI(d)$	$(1 - L)^d w_t = u_{wt}$ for $d \in (0, 1/2)$
11	$FI(1 + d)$	$\Delta w_t = u_{wt}$ with $(1 - L)^d u_{wt} = \epsilon_{wt}$ for $d \in (0, 3/2)$

Mean stationary processes (DGPs 2 and 3) have been used to model the behavior of real exchange rates, unemployment rates, current account, and several great ratios, such as the output-capital ratio and the consumption-income ratio. Unemployment has also been conceived as a nonstationary fractionally integrated process [38]. Some examples can be found in Perron and Vogelsang [39], Wu [40], Wu et al. [41], and D'Adda and Scorcu [42]. Trend Stationarity and $I(2)$ processes (DGPs 4 and 9, resp.) have been used to model growing variables, real and nominal, such as output, consumption and prices; several macro variables have been conceived as DGPs 5, 6, and 7 [9, 43–45]. Variables identified as $I(2)$ processes can also be found in Juselius [46, 47], Haldrup [48], Muscatelli and Spinelli [49], Coenen and Vega [50], and Nielsen [51]. In Table 11 of Appendix B, a few more examples are provided, in which the link between time-series econometrics and important economic issues is acknowledged.

4. Spurious Regression Since the Roaring Twenties

We now begin our survey of the development of the theory of spurious regression. The related literature is vast, for which reason we focus mainly on a limited selection of articles which, in our view, are particularly representative (as noted earlier, Pearson [23], and Yule [22] became aware of several problems concerning the interpretation of the correlation coefficient at the end of the nineteen century, forging by the way the concept of spurious correlation. The debate, in that time, was greatly related with the interpretation of correlation as causality although the seeds of the modern interpretation that prevails in econometrics are also present; further details can be obtained from Aldrich's [25] historical review). Unless otherwise specified, the regression specification for which all asymptotics are presented is hereinafter expression (1). We focus mainly on on the rate of divergence of the relevant *t*-ratios and let aside in most cases the asymptotic distributions that would be obtained had the statistics being correctly normalized. This is so because such distributions do have nuisance parameters that prevent one from making use of these to do correct inference; moreover, practitioners cannot be aware *a priori* of the adequate normalization. However, in some cases [52–54] the asymptotic distributions are important because of the relevance for practitioners

(Valkanov suggests a $T^{1/2}$ normalization of the t-ratio and the simulation of the distribution in the context of Long-Horizon regressions; Sun proposes a consistent t-statistic with a nuisance-free asymptotic distribution, and; Kim, Lee and Newbold precisely emphasize the fact that, even if the t-ratios of their regression do not diverge, inference can not be drawn because of the nuisance parameters).

4.1. Yule's Experiment

Spurious correlation was evidenced by Yule [27] in a computerless Monte Carlo experiment. Shuffling decks of playing cards, Yule obtained independent series of random numbers. In fact, Yule generated independent $I(0)$, $I(1)$, and $I(2)$ series and computed correlation coefficients amongst them. Such correlation coefficients provided correct inference when using $I(0)$ series, but became nonsensical when the order of integration of the variables was higher. With independent $I(0)$ variables the correlation coefficient remained close to zero. The same estimate achieved using $I(1)$ independent variables no longer worked; many times it was close to unity, resulting in what we now refer to as a spurious correlation. If variables were $I(2)$, the most probable outcome was actually a correlation coefficient of close to 1: the spurious phenomenon was even stronger.

4.2. Reappraisal in the Seventies: Spurious Least Squares

Throughout most of the past century, it was commonly recommended to first-difference the series if these seemed to have a trending mechanism. However, not every econometrician was in agreement with such a method because, it was argued, differencing causes losses in the information contained in the original series. A profound reappraisal of this issue began with Granger and Newbold's [1] article, which allowed the spurious regression in Least squares estimators to be identified. The Monte Carlo experiment described above revealed, among other things, that high R^2 and low Durbin-Watson statistics (hereinafter $\mathcal{DW}$) should be considered as a sign of misspecification. They also pointed out, comparing the outcome of simulation with many results in applied econometrics, that problems of misspecification seemed to be widespread. It was proposed that first-differenced series be used, although the authors warned about the risks of catch-all solutions; their results may be considered as the seed of many fruitful extensions in time series econometrics.

4.3. Theory at Last: Asymptotics in Nonsensical Regressions

Phillips [6] proposed the theoretical framework necessary to allow an understanding of Granger and Newbold's earlier results and provided a first insight into the phenomenon of spurious regression. His development set the groundwork for the spurious regression literature in econometrics. Whilst Granger and Newbold used an *i.i.d.* noise in their simulations, Phillips allowed a flexible autocorrelation structure, as well as some degree of heterogeneity. He then proved that, when specification (1) is estimated the following asymptotics in Table 2 are obtained: where $O_p(m)$ denotes the order of magnitude and $\xrightarrow{p}$ means convergence in probability. The most relevant results are that the R^2 does not collapse, and, the t-statistic associated to $\hat{\beta}$ diverges at a rate $T^{1/2}$. This means that the t-statistic will exceed the classical 5%-level critical values (-1.96 and 1.96) as the sample size grows and, with a sufficiently large sample, the null hypothesis of the t-statistic ($\mathcal{H}_0 : \beta = 0$) will be rejected with certainty.

Table 2: Orders in probability: variables y_t and x_t both independently generated by DGP 6.

$\widehat{\beta} = O_p(1)$	$\widehat{\alpha} = O_p(T^{1/2})$	$R^2 = O_p(1)$
$t_{\widehat{\beta}} = O_p(T^{1/2})$	$t_{\widehat{\alpha}} = O_p(T^{1/2})$	$\mathcal{DW} \xrightarrow{p} 0$

Table 3: Orders in probability: variables y_t and x_t both independently generated by DGP 9. Both variables are integrated of the same order.

$\widehat{\beta} = O_p(1)$	$\widehat{\alpha} = O_p(T^{1/2-d})$	$R^2 = O_p(1)$
$t_{\widehat{\beta}} = O_p(T^{1/2})$	$t_{\widehat{\alpha}} = O_p(T^{1/2})$	$\mathcal{DW} = O_p(T^{-2})$

Park and Phillips [55] demonstrated the presence of spurious regression when the independent variables are $I(2)$. Marmol [56] proved that the phenomenon of spurious regression occurs in a more general nonstationary framework; he demonstrated that spurious regression occurs when independent series x_t, y_t are integrated of order d for $d \in \mathcal{N} = \{1, 2, \ldots\}$ (see DGP 9). Empirically relevant DGPs, such as the $I(2)$ process (identified, as mentioned earlier, with uncontrolled inflation) can produce spurious regression. The asymptotic results are somewhat similar to those of Phillips (see Table 3): Although the parameter estimate of the constant term, $\widehat{\alpha}$, diverges at a different rate, the t-statistic associated with $\widehat{\beta}$ diverges at the same rate as in phillips seminal work, that is, $O_p(T^{1/2})$.

One limitation of Marmol's [56] results is that both the dependent and the independent variables share the same order of integration. Banerjee et al. [57], provided Monte Carlo evidence of spurious regression when the order of integration of each variable is different. Marmol [58], later allowed the relevant DGPs for both y and x to be integrated of different orders; $y_t \sim I(d_1)$ and $x_t \sim I(d_2)$ where d_1 and $d_2 \in N = \{1, 2, \ldots\}$. Marmol's [58] results are twofold; either $d_1 > d_2$ or vice versa:

(i) when $d_1 > d_2$ (Table 4):

(ii) when $d_1 < d_2$ (Table 5):

(iii) in both cases $R^2 = O_p(1)$.

Indeed, spurious regression persists in the presence of the so-called unbalanced regressions. The rate of divergence of $t_{\widehat{\beta}}$ remains $T^{1/2}$. De Jong [59] extended the study of spurious regression using independent driftless unit root processes; he used DGP (6) to generate the series and ran the following specification, which operates with logarithmic transformation of both variables—a direction which has proved extremely relevant for empirical purposes: (applying the logarithmic transformation in econometrics is typical when the practitioner wants either to homogenize the variance or to directly obtain estimates of average elasticity amongst variables):

$$\ln(y_t) = \alpha + \beta \ln(x_t) + u_t \tag{4.1}$$

de Jong discovered similar results to previous works, that is, $\widehat{\beta} = O_p(1)$ and $t_{\widehat{\beta}} = O_p(T^{1/2})$.

Entorf [60] made a slight-and yet, fundamental-modification to the Phillips' DGPs by adding a drift (DGP 7) (Entorf's results apply to fixed-effects models where N (the dimension of the cross section) is treated as fixed. The phenomenon of spurious regression in panel data under more general conditions has been further developed; Phillips and Moon [61] provided

Table 4: Orders in probability: variables y_t and x_t both independently generated by DGP 9. Variables are integrated of different orders $d_1 > d_2$.

$\widehat{\beta} = O_p(T^{d_1-d_2})$	$\widehat{\alpha} = O_p(T^{d_1-1/2})$
$t_{\widehat{\beta}} = O_p(T^{1/2})$	$t_{\widehat{\alpha}} = O_p(T^{1/2})$
$\mathcal{DW} = O_p(T^{-1})$ if $d_2 = 1$	$\mathcal{DW} = O_p(T^{-2})$ if $d_2 = 2,3,\ldots$

Table 5: Orders in probability: variables y_t and x_t both independently generated by DGP 9. Variables are integrated of different orders. $d_1 > d_2$.

$\widehat{\beta} = O_p(T^{d_1-d_2})$	$\widehat{\alpha} = O_p(T^{d_1-1/2})$
$t_{\widehat{\beta}} = O_p(T^{1/2})$	$t_{\widehat{\alpha}} = O_p(T^{1/2})$
$\mathcal{DW} = O_p(T^{-1})$ if $d_1 = 1$	$\mathcal{DW} = O_p(T^{-2})$ if $d_1 = 2,3,\ldots$

Table 6: Orders in probability: variables y_t and x_t both independently generated by DGP 7.

$t_{\widehat{\beta}} = O_p(T)$	$t_{\widehat{\alpha}} = O_p(T^{1/2})$
$(1 - R^2) = O_p(T^{-1})$	$\mathcal{DW} = O_p(T^{-1})$

a regression limit theory for non stationary panel data with large numbers of cross section and time series. Kao [62] studied the Least-Squares Dummy Variable estimator (LSDV) where the spurious regression phenomenon is still present for independent nonstationary variables generated by DGPs (6) and (7)). Amongst the most relevant consequences of such drift is the fact that there is not only a stochastic trend but also a deterministic one. In the long run, the deterministic trend dominates the stochastic (see Appendix C). The asymptotic results of estimating (1) using independent variables generated by DGP 7 are given in Table 6: Note that $t_{\widehat{\beta}}$ grows at rate T instead of $T^{1/2}$, contrary to the results presented so far, due to the presence of a deterministic trend.

4.4. Spurious Regression and Long Memory: An Unforgettable Extension

Among the first papers to deal with spurious regression in econometrics (Bhattacharya et al. [5] proved that the R/S test yields spurious inference concerning long memory when the data have a deterministic trend mechanism. See Section 5). using long memory processes are those of Cappuccio and Lubian [63], and Marmol [64]. The authors use the nonstationary fractionally-integrated processes specified in DGP 11. Under these conditions, the asymptotics of an LS regression as specified in expression (1) are given in Table 7 (here the underlying theory is an invariance principle (see Davydov, 1970, Theorem 2, and Sowell, 1990, Theorem 1)).

Note that as most of the previous cases the t-statistic of $\widehat{\beta}$ diverges at rate $T^{1/2}$. The main difference lies in the $\mathcal{DW}$, the rate of divergence of which varies according to the degree of long memory, measured by d. These results can be understood as an argument against the usefulness of the "rule-of-thumb"; a regression was usually considered spurious when $R^2 > DW$. However, when long memory is present in the variables, the regression may well be spurious even if $R^2 < DW$. It is worth mentioning that the fractional integration parameter, d, is the same in both variables. This "shortcoming" is fixed by Tsay and Chung [65], who used a manifold approach. They used two DGPs (No 10 and 11) and then combined these

Table 7: Orders in Probability: variables y_t and x_t both independently generated by DGP 11. The variables are integrated of the same order.

$\widehat{\beta} = O_p(1)$	$t_{\widehat{\beta}} = O_p(T^{1/2})$
$R^2 = O_p(1)$	$\mathcal{DW} = O_p(T^{-(1+2d)})$

in a simple regression. Thus, there are four fractionally-integrated processes, two stationary, [$FI(d_1)$ and $FI(d_2)$ with $d_i \in (0, 1/2)$], and two nonstationary, [$FI(d_i + 1)$ for $d_i \in (1/2, 3/2)$]. The authors then studied several specifications using six DGPs. The first four are used in the estimation of specification (1) whilst the remaining two estimate $y_t = \alpha + \delta t + u_t$:

(1) $y_t \sim FI(1 + d_1)$ and $x_t \sim FI(1 + d_2)$, (the variables may be integrated of order $d \in (1/2, 3/2)$ which generalizes Phillips's [6] results),

(2) $y_t \sim FI(d_1)$ and $x_t \sim FI(d_2)$ where $d_1 + d_2 > 1/2$,

(3) $y_t \sim FI(1 + d_1)$ and $x_t \sim FI(d_2)$ where $d_2 > 1/2$,

(4) $y_t \sim FI(d_1)$ and $x_t \sim FI(1 + d_2)$ where $d_1 > 1/2$,

(5) $y_t \sim FI(1 + d_1)$ where $d_1 > 0$,

(6) $y_t \sim FI(d_1)$ where $d_1 > 0$.

We can summarize the results by stating that $t_{\widehat{\beta}} = O_p(T^r)$ where $0 < r < 1$ which means that the t-ratio always diverges; the divergence rate, r, is, $T^{1/2}$, $T^{d_1 + d_2 - 1/2}$, T^{d_2}, T^{d_1}, $T^{1/2}$, and T^{d_1} for cases (1)–(6) .

Special attention should be given to the Durbin-Watson statistic, which collapses to zero (as usual) except in the second case, where $\mathcal{DW} = O_p(1)$ and hence does not converge to zero as it does elsewhere. This also could be interpreted as an important indication of the need for caution when using the rule-of-thumb, which states that when $R^2 > DW$ there may be spurious regression. As in Cappuccio and Lubian [63], there is evidence that the rule-of-thumb may be a dangerous tool. Another important result appears when we compare model 1 and model 2. The reader may notice that the variables used in the second model are merely the first differences of those used in the first. What is surprising is that the spurious regression persists after the variables have been differenced (the same argument was advanced by Marmol [64]. It should be mentioned that the alternative strategy, "detrending" (removing the trend component in the data by running a regression on time) was also known to provide spurious results [10, 66]). This goes against fifty years of tradition (differencing is used to deal with spurious regression (see Dickey and Pantula [67], e.g.)). Tsay and Chung [65] actually go further by suggesting that the spurious phenomenon is due to the long memory properties of the series and not to the presence of unit roots (see also Cappuccio and Lubian [63]). Sun [68] shows that spurious regression can occur between two stationary generalized fractional processes, as long as their generalized fractional differencing parameters sum up to a value greater than 1/2 and their spectral densities have poles at the same location.

4.5. Spurious Regression With Stationary Series: Size-Matters

Granger et al. [12] show, both through Monte Carlo experiments and theoretically that the spurious regression phenomenon may occur amongst independent stationary series (see also Mikosch and Stărică [69]). Two independent $AR(1)$ series (see DGP 2 with $P_w = 1$) were

run together to estimate regression (1), resulting in a rejection rate of the t-statistic greater than the expected 5%. Granger et al. [12] provide a theoretical proof of why this happens (although the variance of the estimates does not diverge, it may not be unity, depending on the values of the DGPs parameters) (DGP innovations may be drawn from distributions other than normal (Cauchy, Exponential, Laplace and so on), although these remain i.i.d. and independent of each other).

These results are extended to long-span MA processes (see DGP 1). Curiously, the spurious regression does not depend on the sample size, but rather on the MA span parameter, q_w. What is interesting about this result is the fact that it better explains the spurious regression phenomenon in small-sized samples. This would complete the theoretical framework necessary to allow an understanding of the Monte Carlo results of Granger and Newbold [1], Ferson et al. [30].

Granger et al. [12] results differ from all others, the effect being discussed is not an asymptotic phenomenon but rather a size distortion. Size distortions arise because standardizing a test statistic is difficult unless the exact form of the spectral density of the residuals is known. The intuition behind the spurious regression phenomenon is thus, different from the one that underlies all other results.

Mikosch and de Vries [70] provide an alternative theory to explain spurious-type behavior akin to the financial risk measurement literature (Furman and Zitikis [71, Page 7] suggest that "the validity of the CAPM (*Capital Asset Price Model*) is closely related to the linearity of the regression function of the return on asset i given the return on the market portfolio of all assets". Nonlinearities such as mean or variance shifts may induce the spurious regression phenomenon; this implies that the CAPM beta regression estimations should be performed cautiously). In the words of Mikosch and de Vries [70]:

> "Estimators of the coefficients in equations of regression type which involve financial data are often found to vary considerably across different samples. This observation pertains to finance models like the CAPM beta regression, the forward premium and the yield curve regression. In economics, macro models like the monetary model of the foreign exchange rate often yield regression coefficients which significantly deviate from the unitary coefficient on money which is based on the theoretical assumption that money is neutral."

Mikosch and de Vries [70] prove that, when the distribution of the innovations is heavy tailed, that is, when there is a departure from the normality assumption usually made, using standard statistical tools (such as LS, e.g.) can be misleading (the non-normal behavior of financial data has been widely documented (see Fama [72])).

4.6. The Last Newcomer: Trend Stationarity

Hasseler [73] studied the spurious regression phenomenon from a different perspective. He considered the possibility of spurious regression when the variables do have a deterministic trend component (what we have defined earlier as trend stationarity [DGP 4]). Thus, there are two independent nonstationary variables. Hasseler's [73] results, as with those of Kim et al. [52], who worked with the same DGPs, are as in Table 8.

In addition, Kim et al. [52] proved that when one of the deterministic trend components is taken out (i.e., either $\beta_y = 0$ and $\beta_x \neq 0$ or vice versa), the t-ratios converge to centered normal distribution, although its variance is not unity, which may provoke spurious rejection of the null hypothesis that the parameter is zero. That means that running a trend

Table 8: Orders in Probability: variables y_t and x_t both independently generated by DGP 4.

$t_{\widehat{\beta}} = O_p(T^{3/2})$	$\beta = O_p(1)$	$T^2(1 - R^2) \xrightarrow{p} 0$

Table 9: Orders in Probability: variables y_t and x_t both independently generated by DGP 4. The estimated specification is (2).

$\widehat{\alpha} = O_p(1)$	$t_{\widehat{\alpha}} = O_p(T^{1/2})$
$\widehat{\beta} = O_p(T^{-1/2})$	$t_{\widehat{\beta}} = O_p(1)$
$\widehat{\delta} = O_p(1)$	$t_{\widehat{\delta}} = O_p(T^{1/2})$

stationary variable (DGP 4) on a mean stationary variable (DGP 2) or vice versa lessens the "spuriosity" in the regression. The authors also generated independent series with the DGP (4), as did Entorf [60], although they ran a different specification:

$$y_t = \alpha + \beta x_t + \delta t + u_t. \tag{4.2}$$

The asymptotics provided in Kim et al. [74] only concern $\widehat{\beta}$ and $t_{\widehat{\beta}}$. We computed the orders in probability of the other two parameters, given their relevance. We included in Appendix D a guide on how to obtain these asymptotics (Table 9): It should be noted that the t-ratio converges to a normal distribution. However, since the variance is not unity, spurious regression may well be still present.

A relevant extension of Hasseler [73], and Kim et al. [52] was provided by Noriega and Ventosa-Santaulária [75] where (possibly multiple) structural breaks are added to the Trend Stationary DGP, that is, DGP 5. Adding breaks to the specification of the DGPs makes the divergence rate of the t-statistic associated with $\widehat{\beta}$ return to the usual $T^{1/2}$ "norm". They also proved that adding a deterministic trend into the regression specification (see (4.2) does not prevent the phenomenon of spurious regression; $t_{\widehat{\delta}}$ remains $O_p(T^{1/2})$. In an effort to unify the related literature, Noriega and Ventosa-Santaulária [76] filled a number of gaps until then unaddressed. They studied several previously overlooked combinations of DGPs and summarized most of the others' results. DGPs 2-9 may indistinctly generate x and/or y (obviously not every combination was carried out by Noriega and Ventosa-Santaulária [76]; besides those presented previously, the following combinations where made by Hasseler [77, 78]: DGP 6-DGP 2, DGP 2-DGP 3, DGP 3-DGP 2, and, DGP 3-DGP 3). In fact, amongst the new possible combinations, the divergence rate of $t_{\widehat{\beta}}$ usually remains $T^{1/2}$. The combination of a Trend Stationary and an $I(1)$ plus drift stands out since the divergence rate of the t-ratio is T rather than $T^{1/2}$ as in Kim et al. [52], although this result should have been anticipated given the asymptotic dominance of the deterministic trend over the stochastic. This combination is thus clearly linked to Entorf's [60] results.

Granger et al. [12] showed that the use of an Heteroskedasticity and Autocovariance covariance matrix diminishes size distortions in some cases but their Monte Carlo evidence also showed that this is true only when the sample size is large (greater than 500). The use of HAC standard errors (see Newey and West [79]). when the DGPs are stationary or have a deterministic trend is not so obvious. We performed a simple simulation to support this

Table 10: t-ratio rejection rates using a 1.96 critical value (5% level) for a standard normal distribution: spurious regression using LS and HAC standard errors.

DGPs		LS standard errors				HAC standard errors			
					Sample size				
x_i	y_i	50	100	200	500	50	100	200	500
4	4	0.21	0.99	1.00	1.00	0.27	0.99	1.00	1.00
5	5	0.20	0.99	1.00	1.00	0.27	0.99	1.00	1.00
6	4	0.45	0.80	0.90	0.95	0.46	0.71	0.79	0.86
6	5	0.65	0.84	0.91	0.95	0.62	0.74	0.80	0.86
6	6	0.68	0.76	0.83	0.90	0.57	0.61	0.67	0.77

view (see Table 10). The variables x and y are independently generated either by DGP (4), (5), or (6). Lag selection was done as in Granger et al. [12], that is, using the formula: $l = $ integer$[4(T/100)^{1/4}]$. Simulation results show that size distortions are less severe when HAC is used, but remain extremely high (perhaps results would improve further if the data could follow a pre-whitening procedure).

4.7. Next of Kin: Statistical Tests, Long Horizon, and Instrumental Variables

Not all work on the spurious phenomenon has been monopolized by the LS simple estimator. Stochastic trending mechanisms have been used in most of the studies, although some exceptions are presented at the end of this section. Equally relevant are the Instrumental Variables (IV) estimates, which have also been analyzed [80].

Giles [81] studied two important residual tests: the Jarque-Bera test [hereinafter $\mathcal{JB}$; see Jarque and Bera [82]] and the Breusch-Pagan-Godfrey test [83, 84], used to test for normality and autocorrelation/heteroskedasticity evidence in the residuals of an LS regression, respectively; he used independent driftless random walks (DGP 6) and proved that both statistics diverge at rate T, that is:

(i) Jarque-Bera test: $\mathcal{JB} = O_p(T)$

(ii) Breusch-Pagan-Godfrey test: $(T \cdot R^2) = O_p(T)$.

The null hypothesis of normality or serial independence/homoskedasticity will be eventually rejected even if it is correct, given a large enough sample (these residuals are neither normal nor serially independent or homoskedastic; hence, the statistical tests provide appropriate inference by rejecting the null). Furthermore, Ventosa-Santaulária and Vera-Valdés [85] studied the behavior of the classical Granger-Causality test (see Granger [86]), hereinafter GC. It is proved that the classical GC test fails to accept the null hypothesis of no GC between independent broken-trend (DGP 5) or broken-mean (DGP 3) processes whether the former series are differenced or not.

Kim et al. [87] study several important nonlinearity tests: the RESET test [88], the McLeod and Li test [89], the Keenan test [90], the Neural Network test [91], White's Information test [92], and the Hamilton test [93]. All these tests were studied by way of finite-sample experiments using independent driftless random walks (see DGP 6). They also presented the asymptotics of the RESET and Keenan tests; both test-statistics are $O_p(T)$, so the null of linearity is eventually rejected as the sample size grows. In the same vein, Noriega

and Ventosa-Santaulária [94] proved that, when the variables are generated independently by any combination of DGPs (3), (5), (7), and (8), the Engle and Granger cointegration-test spuriously rejects the null of no cointegration (of course, there is an important exception: if both variables are independent $I(1)$ processes, the Engle and Granger test tends to accepts the null of no-cointegration). in an indeterminate number of cases since the relevant t-statistic is $O_p(T^{1/2})$ (the t-statistic may diverge toward $+\infty$ or $-\infty$; the sign depends on the unknown values of the DGPs; Monte Carlo evidence is presented to assert that in many cases, such sign is negative).

Spurious regression has also been identified in the context of Long-Horizon (henceforth LH) regressions, which are used in situations where previous "short-term" (that is, using the variables in levels, without doing any kind of temporal aggregation). studies have failed. In fact, Valkanov's [95] insight is that rolling summations of $I(0)$ variables, that is LH variables, behave asymptotically as $I(1)$ series. Hence, the theory of spurious regression between independent $I(1)$ variables, discussed earlier, should give readers the intuition. Late eighties' studies provided interesting results in economics and finance. As asserted by Valkanov [95]:

> "The results [...] are based on long-horizon variables, where a long-horizon variable is obtained as a rolling sum of the original series. It is heuristically argued that long-run regressions produce more accurate results by strengthening the signal coming from the data, while eliminating the noise."

More precisely, if the specification to be estimated is (1), a Long-Horizon "reinterpretation" is carried out with the building of partially aggregated variables. As usual, let, $w = x, y$, then $w_t^k = \sum_{j=0}^{k-1} w_{t+j}$. The LH regression specification can usually be one of the following:

$$y_t^k = \alpha + \beta x_t^k + u_t,$$
$$y_t^k = \alpha + \beta x_{t-k}^k + u_t, \tag{4.3}$$
$$y_t^k = \alpha + \beta x_t + u_t.$$

Such specifications are used mostly in the estimation of the equity/dividend relationship and to test both the neutrality of money or the Fisher effect. Based upon the previous specifications, Valkanov [95], and Moon et al. [54] proved that this regression strategy also presents the spurious regression phenomenon. To do this, they let time overlap in the summations as a fixed fraction of the sample size, $k = [\lambda \cdot T]$ where $\lambda \in [0, 1]$. Valkanov [95] then defines the following DGP:

$$y_t = \beta x_t + u_{yt}, \tag{4.4}$$

where $(1 - L)\phi(L)x_t = u_{xt}$. The variable x_t follows an autoregressive process whose highest root is unity whilst the rest, represented in the polynomial $\phi(L)$, is invertible. Let $\omega = (u_{yt}, u_{xt})'$ be defined as a martingale difference sequence with $E(\omega_t \omega_t' \mid \omega_{t-1} \cdots) = [\sigma_{11}^2 \ \sigma_{12}; \sigma_{21} \ \sigma_{22}^2]$ and finite fourth moment. We could argue that, when $\beta = 0$, the t-ratio associated with its estimate should be small enough for the null hypothesis to be accepted.

As in previous results, that does not happen; $t_{\hat{\beta}} = O_p(T^{1/2})$. Valkanov [95], and Moon et al. [54] suggest the use of a rescaled t-ratio, $t_{\hat{\beta}}/T^{1/2}$. Although the limiting distribution of such t-statistic's is neither normal nor pivotal, it can be easily simulated and hence used. Lee [96] extends the results for fractionally integrated processes and finds that the t-ratio associated with $\hat{\beta}$ in (3) diverges: $t_{\hat{\beta}} = O_p(T^{1/2})$.

Ventosa-Santaulária [80] studies the asymptotics of Instrumental Variables (IV) using DGPs (5) and (7), not only for x and y, but also for a spurious and independent instrument. It is shown that the t-ratio still diverges at rate $T^{1/2}$ which confirms the Monte Carlo simulations carried out by Leybourne and Newbold [97]. This result can be also seen as a complement to those presented by Phillips and Hansen [98], and Hansen and Phillips [99], in which the use of spurious instruments is proposed to improve the estimates whenever strong endogeneity is present between x and the residual term in cointegrated relationships.

Sun [53] developed a convergent t-statistic for correcting the phenomenon of spurious regression. The new t-ratio is based on an estimate of the parameter variance made in the same manner as HAC standard errors. This t-statistic converges to a non-degenerate limiting distribution for many cases of spurious regression. He considered the regression between two independent nonstationary $I(d)$ processes with $d > 1/2$ as well as the regression between an independent nonstationary $I(d)$ process and a linear trend (see (4.5)):

$$y_t = \alpha + \delta t + u_t. \tag{4.5}$$

In previous studies, the presence of unit roots in the series generally produced a $T^{1/2}$ divergent t-ratio. To avoid such divergence, Sun proposes a rescaling of the parameter estimate using a new standard error. The new Standard Error is computed in the same way HAC is. The main difference is that HAC estimates usually require a bandwidth or truncation lag. Sun suggests using the entire sample length:

$$\hat{\sigma}_{\hat{\beta}}^2 = \left(\sum_{t=1}^{T} (x_t - \overline{x})^2 \right)^{-1} \cdot T\hat{\Omega} \cdot \left(\sum_{t=1}^{T} (x_t - \overline{x})^2 \right)^{-1}, \tag{4.6}$$

where

$$\hat{\Omega} = \sum_{j=-T+1}^{T-1} \kappa \left(\frac{j}{T} \right) \hat{\Gamma}(j),$$

$$\hat{\Gamma}(j) = \begin{cases} \dfrac{1}{T} \displaystyle\sum_{t=1}^{T-j} (x_{t+j} - \overline{x}) \, \hat{u}_{t+j} \hat{u}_t \, (x_t - \overline{x}) & \text{for } j \geq 0, \\[4mm] \dfrac{1}{T} \displaystyle\sum_{t=-j+1}^{T} (x_{t+j} - \overline{x}) \, \hat{u}_{t+j} \hat{u}_t \, (x_t - \overline{x}) & \text{for } j < 0, \end{cases} \tag{4.7}$$

where κ is a kernel function that belongs to a class that ensures positive definitiveness. Whether $x_t \sim I(d)$ or $x_t = t$, well-defined asymptotic distributions for the t-ratio are provided.

5. What to do if One Fears Spurious Regression

The spurious regression phenomenon pervades many subfields in time series analysis. It might be controlled by using correctly scaled t-ratios, as suggested earlier, but having a clear idea of what DGP best emulates the properties of the series (this could be labeled as *DGPi-fication*) would therefore be necessary (examples of this can be: (i) evidence of UR must be obtained before a cointegration analysis is undertaken; (ii) the nature of the trending mechanism should be identified prior to the application of a transformation intended to render stationary the series, and; (iii) a test such as Robinson's [100] could be undertaken in order to identify long-memory behavior). This can be achieved by means of applying a battery of tests to our series. Such an approach is not exempted of failures. Many statistical tests are known to yield spurious evidence under specific circumstances (see the previous subsection). Nevertheless, pretesting the series remains an adequate strategy and allows the practitioner to be aware of the potencial difficulties he could face. In this section we include a short—and incomplete—list of tests that are employed in the *DGPi-fication* of the series:

(1) Drawing inference concerning the nonstationarity of the series can be done by means of Dickey-Fuller-type (DF) tests (see Dickey and Fuller [8, 101]). The original DF test distinguishes between the null hypotheses of UR (DGPs 6 and 7) and the alternatives of stationarity (DGPs 1, 2 and 4). Other well-known UR tests are: (i) the KPSS test [102]; (ii) the GLS-detrended DF test [103]; (iii) the Phillips-Perron test [104], and; (iv) the Ng and Perron test [105, 106].

(2) The UR tests previously mentioned provide severely biased results under the hypothesis of trend-stationarity in the presence of structural breaks: DF-type tests (i) over-accept the null hypothesis of unit root when there is a trend/level break in the—trend—stationary process [9, 107–113], and; (ii) over-reject the null hypothesis when there is a trend/level/variance break in the unit root process [114–117]. Several alternatives are available. Perron [9] suggested the use of a DF-type test with (point,level and trend) breaks specified in the auxiliary regression (see DGP 5); the break dates must be decided by the practitioner. Zivot and Andrews [118] also proposed to modify the DF test in the same direction than Perron, only they allowed the break date to be endogenously specified; their test allows for a single break (see Lumsdaine and Papell [44] for an extension that allows for two breaks) under the alternative hypothesis (DGP 5) and rules out the possibility of a break under the null hypothesis of UR (DGP 8); Carrion-i-Silvestre and Sansó [119] proposed a test where a break under the null hypothesis is taken into account.

(3) Bai and Persson [120] proposed a test to distinguish between DGPs 2 and 4 and DGPs 3 and 5. The test presupposes that the trending mechanism is exclusively deterministic.

(4) Long-range dependence: many testing procedures are also available. R/S-type test (see Hurst [32], Mandelbrot and Taqqu [121], and Lo [122]) are commonly used to identify Long Memory (LM) against stationarity. However, Bhattacharya et al. [5] proved that the classical R/S test may provide evidence of LM even if the series is

stationary when the latter contains a trending mechanism. Moreover, Mikosch and Stărică [123], and Mikosch and Stărică [69] proved that the sample autocorrelation function (sample ACF) can also be a misleading statistical tool when used to identify LM; stationary series that include a non-linear component might yield a sample ACF usually attributed to LM processes (see also Teverovsky and Taqqu [124]). Several Short memory processes may thus seem to behave as LM processes. This phenomenon can be labeled as spurious long memory [69].

(5) Many other tests have been proposed to identify LM while they control for possible non-linearities (structural breaks in the mean or the variance of the series). See Liu et al. [125], Robinson [100], Lobato and Robinson [126], Giraitis et al. [127, 128], Berkes et al. [129], Zhang et al. [130], Aue et al. [131], and Jach and Kokoszka [132].

6. Concluding Remarks

Spurious regression can arise wherever a trending mechanism is present in the data. Even some stationary-autocorrelated processes cause spurious results.

Applied macroeconomists and financial experts have been steadily incorporating technical advances in the analysis of the spurious regression, a phenomenon identified for many empirically-relevant Data Generating Processes. These include stationary processes with AR (or long MA) structure and/or level breaks; random walks (with or without drifts), trend stationarity (with possible level and trend breaks), long-memory processes (whether stationary or not) and so on. These processes have been associated with unemployment rates, price levels, real exchange rates, monetary aggregates, gross domestic product and various financial variables. The use of Least Squares with such variables entails a high risk of obtaining a spurious relationship.

Differencing the series may not always prevent spurious estimates; nor should the $R^2 > DW$ rule-of-thumb be seen as an adequate rule to identify a spurious regression. Cointegration analysis appears to better prevent non-sensical statistical relationships although, one should bear in mind Phillips's [20] and further study the statistical relationship at hand. Out-of-sample forecasting evaluation could be an option. Most macroeconomic variables are either nonstationary or very persistent. Pre-testing the variables in order to identify the nature of the trending mechanism arises as the golden rule to avoid non-sense regression. Once the DGP is correctly identified, spurious regression is "easier" to deal with.

Attaining a clear understanding of any problem is the first step toward finding its solution.

Appendices

A. Spurious Regression Using Independent UR Processes

The theoretical explanation of the spurious regression phenomenon documented in Granger and Newbold [1] was provided by Phillips [6]. We present here a simple version of Phillips's [6] results. Assume two independent nonstationary unit root processes:

$$x_t = x_{t-1} + u_{xt},$$
$$y_t = y_{t-1} + u_{yt}, \tag{A.1}$$

where, for simplicity, $u_{zt} \sim \mathcal{N}(0, \sigma_z^2)$; $u_{xt} \perp u_y t$. By solving recursively both equations, we obtain:

$$
x_t = X_0 + \underbrace{\sum_{i=1}^{t-1} u_{xt-i}}_{\xi_{xt-1}},
$$

$$
y_t = Y_0 + \underbrace{\sum_{i=1}^{t-1} u_{yt-i}}_{\xi_{yt-1}},
$$

(A.2)

where X_0 and Y_0 are initial conditions. It can be proved (see, e.g., Hamilton [133, Pages 486 and 548]) that:

$$
T^{-3/2} \sum_{t=1}^{T} \xi_{zt-1} \xrightarrow{d} \sigma_z \int_0^1 \omega_z(r) \, dr,
$$

$$
T^{-2} \sum_{t=1}^{T} \xi_{zt-1}^2 \xrightarrow{d} \sigma_z^2 \int_0^1 [\omega_z(r)]^2 dr,
$$

(A.3)

$$
T^{-2} \sum_{t=1}^{T} \xi_{xt-1} \xi_{yt-1} \xrightarrow{d} \sigma_x \sigma_y \int_0^1 \omega_x(r) \omega_y(r) \, dr,
$$

where $\omega_z(1)$ is a standard brownian motion and $\xrightarrow{d}$ denotes convergence in law. Therefore (assume that $u_{z0} = 0$):

$$
\sum z_t = Z_0 T + \underbrace{\sum_{t=1}^{T} \xi_{zt-1}}_{O_p(T^{3/2})},
$$

$$
\sum z_t^2 = Z_0^2 T + \underbrace{\sum_{t=1}^{T} \xi_{zt-1}^2}_{O_p(T^2)} + \sum_{t=1}^{T} \xi_{zt-1},
$$

(A.4)

$$
\sum x_t y_t = X_0 Y_0 T + X_0 \sum_{t=1}^{T} \xi_{yt-1} + Y_0 \sum_{t=1}^{T} \xi_{xt-1} + \underbrace{\sum_{t=1}^{T} \xi_{xt-1} \xi_{yt-1}}_{O_p(T^2)}.
$$

The LS formula of $\widehat{\beta}$ is:

$$
\widehat{\beta} = \frac{T \sum x_t y_t - \sum x_t \sum y_t}{T \sum x_t^2 - \left(\sum x_t\right)^2}.
$$

(A.5)

Table 11: Empirical applications in macroeconomics and finance.

Subject or field	Selection of articles
GDP and GNP trending, growth, convergence and Business cycles, Output-Capital Ratio, exports, unemployment,	[7, 9, 38, 42, 43, 45, 51, 118, 134–143]
Quantitative theory and Money neutrality, Demand of Money,	[4, 49, 50, 54, 95, 144–146]
Purchasing Power Parity, Exchange Rates, Real interest Rate, Taylor Rule, Inflation, Persistence, Stock Prices, CAPM, Forward premium , Yied curve, Financial risk, Premium calculation principles (pcp's),	[39, 46, 47, 71, 147–165]
Phillip's Curve, Fisher , Balassa-Samuelson Effect,	[166–176]

Replacing the sums that appear in this formula with the above asymptotic expressions and letting $T \to \infty$ yields:

$$
\hat{\beta} = \frac{\sigma_y}{\sigma_x} \frac{\int_0^1 \omega_x(r)\,\omega_y(r)\,dr - \int_0^1 \omega_x(r)\,dr \int_0^1 \omega_y(r)\,dr}{\int_0^1 [\omega_x(r)]^2 dr - \left(\int_0^1 \omega_x(r)\,dr\right)^2}
$$

$$
= O_p(1). \tag{A.6}
$$

The same can be done with the variance of the regression as well as with the t-ratio associated to $\hat{\beta}$:

$$
\hat{\sigma}^2 \xrightarrow{d} T^{-1} \left[\sum y_t^2 + \hat{\alpha}^2 T + \hat{\beta}^2 \sum x_t^2 - 2\hat{\alpha} \sum y_t - 2\hat{\beta} \sum x_t y_t + 2\hat{\alpha}\hat{\beta} \sum x_t \right]
$$

$$
= O_p(T), \tag{A.7}
$$

$$
t_{\hat{\beta}} = O_p\left(T^{1/2}\right).
$$

Note that the t-ratio diverges as the sample size grows. The usual critical values to test the null hypothesis $\beta = 0$ are ± 1.96 (level: 5%) so it is straight forward to see that the null hypothesis will always be rejected for a sample sufficiently large: this is a simple example of spurious regression.

B. Empirical Work and DGPs

Table 11 explains empirical applications in macroeconomics and finance.

C. Dominance of the Deterministic Trend Over the Stochastic Trend

In order to explain the dominance of the deterministic trend over the stochastic trend we may study the sums of a UR with drift. Assume x_t is generated by DGP 7; $x_t = \mu_x + x_{t-1} + u_{xt}$. By solving recursively this we get:

$$x_t = \underbrace{X_0}_{1} + \underbrace{\mu_x t}_{2} + \underbrace{\xi_{xt-1}}_{3}, \tag{C.1}$$

where the first term (1) is an initial condition, the second term (2) represents the deterministic trend, and the third term (3) accounts for the stochastic trend; note that $\xi_{xt} = \sum_{i=1}^{t} u_{xt-i}$. Then (all sums run from 1 to T),

$$\sum x_{t-1} = \underbrace{X_0 T}_{O_p(T)} + \mu_x \underbrace{\sum (t-1)}_{O_p(T^2)} + \underbrace{\sum \xi_{xt-1}}_{O_p(T^{3/2})}, \tag{C.2}$$

$$\sum x_{t-1}^2 = X_0^2 T + \underbrace{\mu_x^2 \sum (t-1)^2}_{O_p(T^3)} + \underbrace{\sum \xi_{xt-1}^2}_{O_p(T^2)} + 2X_0 \mu_x \sum (t-1)$$

$$+ 2X_0 \sum \xi_{xt-1} + \underbrace{\sum (t-1)\xi_{xt-1}}_{O_p(T^{5/2})}. \tag{C.3}$$

The orders in probability of the underbraced sums can be found in Hamilton [133, chapters 16 and 17]. Note that the leading term in expressions (C.2) and (C.3) is always the deterministic component (T^2 in $\sum x_{t-1}$ and T^3 in $\sum x_{t-1}^2$). This is why it is said that the time trend asymptotically dominates the stochastic components.

D. Asymptotics of LS Estimates of Specification 2

We present a guide as to how to obtain the order in probability of the estimates and their associated t-ratios using LS where the variables y and x are generated by DGP 4 and specification 2 is estimated. Proof of such was provided with the aid of *Mathematica 4.1* software. We use the classical LS. Let $\widehat{\Theta} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\delta})'$:

$$\widehat{\Theta} = (X'X)^{-1} X'Y,$$

$$\mathrm{Var}(\widehat{\Theta}) = \widehat{\sigma}^2 (X'X)^{-1},$$

$$t_{\widehat{\alpha}} = \frac{\widehat{\alpha}}{\sqrt{\widehat{\sigma}_{\widehat{\alpha}}^2}},$$

$$t_{\widehat{\beta}} = \frac{\widehat{\beta}}{\sqrt{\widehat{\sigma}^2_{\widehat{\beta}}}},$$

$$t_{\widehat{\delta}} = \frac{\widehat{\delta}}{\sqrt{\widehat{\sigma}^2_{\widehat{\delta}}}},$$

(D.1)

where

$$X'X = \begin{bmatrix} T & \sum x_t & \sum t \\ \sum x_t & \sum x_t^2 & \sum x_t t \\ \sum t & \sum x_t t & \sum t^2 \end{bmatrix}, \qquad Z'Y = \begin{bmatrix} \sum y_t \\ \sum x_t y_t \\ \sum y_t t \end{bmatrix},$$

$$\widehat{\sigma}^2 = T^{-1} \left[\sum y_t^2 + \widehat{\alpha}^2 T + \widehat{\beta}^2 \sum x_t^2 + \widehat{\delta}^2 \sum t^2 - 2\widehat{\alpha} \sum y_t - 2\widehat{\beta} \sum x_t y_t \right.$$

$$\left. - 2\widehat{\delta} \sum y_t t + 2\widehat{\alpha}\widehat{\beta} \sum x_t + 2\widehat{\alpha}\widehat{\delta} \sum t + 2\widehat{\beta}\widehat{\delta} \sum x_t t \right].$$

(D.2)

We will now describe the process involved in establishing the aforementioned proof. $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\delta}$ and their corresponding t-ratios are functions of the following expressions (unless indicated otherwise, all sums run from $t = 1$ to T). Let $w = y, x$:

$$\sum w_t = \mu_w T + \beta_w \sum t + \underbrace{\sum u_{wt}}_{O_p(T^{1/2})},$$

$$\sum w_t^2 = \mu_w^2 T + \beta_w^2 \sum t^2 + \underbrace{\sum u_{wt}^2}_{O_p(T)} + 2\mu_w \beta_w \sum t$$

$$+ 2\mu_w \sum u_{wt} + 2\beta_w \underbrace{\sum u_{wt} t}_{O_p(T^{3/2})},$$

$$\sum w_t t = \mu_w \sum t + \beta_w^2 \sum t^2 + \sum u_{wt} t,$$

$$\sum x_t y_t = \mu_y \sum x_t + \beta_y \sum x_t t + \sum x_t u_{yt},$$

$$\sum x_t u_{yt} = \mu_x \sum u_{yt} + \beta_x \sum u_{yt} t + \underbrace{\sum u_{xt} u_{yt}}_{O_p(T^{1/2})},$$

(D.3)

where

$$\sum t = \frac{1}{2} \left(T^2 + T \right),$$

$$\sum t^2 = \frac{1}{6} \left(2T^3 + 3T^2 + T \right).$$

(D.4)

The orders in convergence of the underbraced expressions can be found in Hamilton [133, Chapter 16]. We can fill the previously-cited matrices and then compute the LS parameter estimates and the *t*-statistic associated with each one. The asymptotics are computed by the program (available at the following URL: http://www.ventosa-santaularia.com/JPSappD.pdf).

References

[1] C. W. J. Granger and P. Newbold, "Spurious regressions in econometrics," *Journal of Econometrics*, vol. 2, no. 2, pp. 111–120, 1974.

[2] J. Davidson, D. Hendry, F. Srba, and S. Yeo, "Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom," *Economic Journal*, vol. 88, pp. 661–692, 1978.

[3] D. Hendry and G. Mizon, "Serial correlation as convenience simplification, not a nuissance: a comment on a study of the demand for money by the bank of England," *Economic Journal*, vol. 88, pp. 549–563, 1978.

[4] C. I. Plosser and G. W. Schwert, "Money, income, and sunspots: measuring economic relationships and the effects of differencing," *Journal of Monetary Economics*, vol. 4, no. 4, pp. 637–660, 1978.

[5] R. N. Bhattacharya, V. K. Gupta, and Ed Waymire, "The Hurst effect under trends," *Journal of Applied Probability*, vol. 20, no. 3, pp. 649–662, 1983.

[6] P. C. B. Phillips, "Understanding spurious regressions in econometrics," *Journal of Econometrics*, vol. 33, no. 3, pp. 311–340, 1986.

[7] C. R. Nelson and C. R. Plosser, "Trends and random walks in macroeconmic time series: some evidence and implications," *Journal of Monetary Economics*, vol. 10, no. 2, pp. 139–162, 1982.

[8] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366, part 1, pp. 427–431, 1979.

[9] P. Perron, "The great crash, the oil price shock and the unit root hypothesis," *Econometrica*, vol. 57, pp. 1361–1401, 1989.

[10] S. N. Durlauf and P. C. B. Phillips, "Trends versus random walks in time series analysis," *Econometrica*, vol. 56, no. 6, pp. 1333–1354, 1988.

[11] P. C. B. Phillips, "Trending time series and macroeconomic activity: some present and future challenges," *Journal of Econometrics*, vol. 100, no. 1, pp. 21–27, 2001.

[12] C. W. J. Granger, N. Hyung, and Y. Jeon, "Spurious regressions with stationary series," *Applied Economics*, vol. 33, no. 7, pp. 899–904, 2001.

[13] G. Sargan, *Wages and Prices in United Kingdom: A Study in Econometric Methodology*, Buttewoth, London, UK, 1964.

[14] D. Hendry and G. Anderson, "Testing dynamic specification in small simultaneous systems: an application to a model of building society behaviour in the United Kingdom," *Frontiers in Quantitative Economics*, vol. 3, pp. 361–383, 1977.

[15] D. Hendry, A. Pagan, and J. Sargan, "Dynamic specification," *Handbook of Econometrics*, vol. 2, pp. 1023–1100, 1984.

[16] C. W. J. Granger, "Some properties of time series data and their use in econometric model specification," *Journal of Econometrics*, vol. 16, no. 1, pp. 121–130, 1981.

[17] C. W. J. Granger and A. A. Weiss, "Time series analysis of error-correction models," in *Studies in Econometrics, Time Series, and Multivariate Statistics*, pp. 255–278, Academic Press, New York, NY, USA, 1983.

[18] R. F. Engle and C. W. J. Granger, "Co-integration and error correction: representation, estimation, and testing," *Econometrica*, vol. 55, no. 2, pp. 251–276, 1987.

[19] P. C. B. Phillips, "New tools for understanding spurious regressions," *Econometrica*, vol. 66, no. 6, pp. 1299–1325, 1998.

[20] P. C. B. Phillips, "Laws and limits of econometrics," *Economic Journal*, vol. 113, no. 486, pp. C26–C52, 2003.

[21] A. Legendre, "Nouvelles méthodes pour la détermination des orbites des comètes," F. Didot, 1805.

[22] G. Yule, "On the theory of correlation," *Journal of the Royal Statistical Society*, vol. 60, no. 4, pp. 812–854, 1897.

[23] K. Pearson, "On a form of spurious correlation which may arise when indices are used in the measurement of organs," *Proceedings of the Royal Society of London*, vol. 60, pp. 489–498, 1897.

[24] D. F. Hendry and M. S. Morgan, *The Foundations of Econometric Analysis*, Cambridge University Press, Cambridge, UK, 1995.

[25] J. Aldrich, "Correlations genuine and spurious in Pearson and Yule," *Statistical Science*, vol. 10, no. 4, pp. 364–376, 1995.

[26] M. S. Morgan, *The History of Econometric Ideas*, Historical Perspectives on Modern Economics, Cambridge University Press, Cambridge, UK, 1990.

[27] G. Yule, "Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series," *Journal of the Royal Statistical Society*, vol. 89, no. 1, pp. 1–63, 1926.

[28] D. Hendry, "Econometrics-alchemy or science?" *Economica*, vol. 47, no. 188, pp. 387–406, 1980.

[29] W. Jevons, "Commercial crises and sunspots," in *Investigations in Currencies and Finance*, pp. 221–243, MacMillan, London, UK, 1884.

[30] W. Ferson, S. Sarkissian, and T. Simin, "Is stock return predictability spurious?" *Journal of Investment Management*, vol. 1, no. 3, pp. 1–10, 2003.

[31] J. H. Stock, "Asymptotic properties of least squares estimators of cointegrating vectors," *Econometrica*, vol. 55, no. 5, pp. 1035–1056, 1987.

[32] H. Hurst, "Long term storage capacity reservoirs," *Transcations of The American Society of Civil Engineers*, vol. 116, pp. 770–799, 1951.

[33] K. W. Hipel and A. I. McLeod, "Preservation of the rescaled adjusted range parts 1, 2 and 3," *Water Resources Research*, vol. 14, pp. 491–518, 1978.

[34] I. Adelman, "Long cycles: fact or artifact," *American Economic Review*, vol. 55, pp. 444–463, 1965.

[35] C. W. J. Granger, "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, vol. 14, no. 2, pp. 227–238, 1980.

[36] C. W. J. Granger and R. Joyeux, "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis*, vol. 1, no. 1, pp. 15–29, 1980.

[37] J. R. M. Hosking, "Fractional differencing," *Biometrika*, vol. 68, no. 1, pp. 165–176, 1981.

[38] M. A. Ariño and F. Marmol, "A permanent-transitory decomposition for ARFIMA processes," *Journal of Statistical Planning and Inference*, vol. 124, no. 1, pp. 87–97, 2004.

[39] P. Perron and T. Vogelsang, "Nonstationarity and level shifts with an application to purchasing power parity," *Journal of Business and Economic Statistics*, vol. 10, no. 3, pp. 301–320, 1992.

[40] J.-L. Wu, "Mean reversion of the current account: evidence from the panel data unit-root test," *Economics Letters*, vol. 66, no. 2, pp. 215–222, 2000.

[41] J.-L. Wu, S.-L. Chen, and H.-Y. Lee, "Are current account deficits sustainable? Evidence from panel cointegration," *Economics Letters*, vol. 72, no. 2, pp. 219–224, 2001.

[42] C. D'Adda and A. E. Scorcu, "On the time stability of the output-capital ratio," *Economic Modelling*, vol. 20, no. 6, pp. 1175–1189, 2003.

[43] P. Perron, "Further evidence on breaking trend functions in macroeconomic variables," *Journal of Econometrics*, vol. 80, no. 2, pp. 355–385, 1997.

[44] R. L. Lumsdaine and D. H. Papell, "Multiple trend breaks and the unit-root hypothesis," *The Review of Economics and Statistics*, vol. 79, no. 2, pp. 212–218, 1997.

[45] A. Mehl, "Unit root tests with double trend breaks and the 1990s recession in Japan," *Japan and the World Economy*, vol. 12, no. 4, pp. 363–379, 2000.

[46] K. Juselius, "A structured VAR under changing monetary policy," *Journal of Business and Economics Statistics*, vol. 16, pp. 400–412, 1996.

[47] K. Juselius, "Price convergence in the medium and long run: an I (2) analysis of six price indices," in *Cointegration, Causality and Forecasting: A Festschrift to Clive Granger*, Oxford University Press, Oxford, UK, 1999.

[48] N. Haldrup, "An econometric analysis of 1(2) variables," *Journal of Economic Surveys*, vol. 12, no. 5, pp. 645–650, 1998.

[49] V. A. Muscatelli and F. Spinelli, "The long-run stability of the demand for money: Italy 1861–1996," *Journal of Monetary Economics*, vol. 45, no. 3, pp. 717–739, 2000.

[50] G. Coenen and J.-L. Vega, "The demand for M3 in the Euro area," *Journal of Applied Econometrics*, vol. 16, no. 6, pp. 727–748, 2001.

[51] H. B. Nielsen, "An I(2) cointegration analysis of price and quantity formation in Danish manufactured exports," *Oxford Bulletin of Economics & Statistics*, vol. 64, no. 5, pp. 449–472, 2002.

[52] T.-H. Kim, Y.-S. Lee, and P. Newbold, "Spurious regressions with stationary processes around linear trends," *Economics Letters*, vol. 83, no. 2, pp. 257–262, 2004.

[53] Y. Sun, "A convergent *t*-statistic in spurious regressions," *Econometric Theory*, vol. 20, no. 5, pp. 943–962, 2004.

[54] R. Moon, A. Rubia, and R. Valkanov, "Long-horizon regressions when the predictor is slowly varying," *Journal of Business and Economic Statistics*, 2006.

[55] J. Y. Park and P. C. B. Phillips, "Statistical inference in regressions with integrated processes. II," *Econometric Theory*, vol. 5, no. 1, pp. 95–131, 1989.

[56] F. Marmol, "Spurious regressions between I(d) processes," *Journal of Time Series Analysis*, vol. 16, no. 3, pp. 313–321, 1995.

[57] A. Banerjee, J. Dolado, J. Galbraith, and D. Hendry, *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press, Oxford, UK, 1993.

[58] F. Marmol, "Nonsense regressions between integrated processes of different orders," *Oxford Bulletin of Economics & Statistics*, vol. 58, no. 2, pp. 525–534, 1996.

[59] R. M. de Jong, "Logarithmic spurious regressions," *Economics Letters*, vol. 81, no. 1, pp. 13–21, 2003.

[60] H. Entorf, "Random walks with drifts: nonsense regression and spurious fixed-effect estimation," *Journal of Econometrics*, vol. 80, no. 2, pp. 287–296, 1997.

[61] P. C. B. Phillips and H. R. Moon, "Linear regression limit theory for nonstationary panel data," *Econometrica*, vol. 67, no. 5, pp. 1057–1111, 1999.

[62] C. Kao, "Spurious regression and residual-based tests for cointegration in panel data," *Journal of Econometrics*, vol. 90, no. 1, pp. 1–44, 1999.

[63] N. Cappuccio and D. Lubian, "Spurious regressions between I(1) processes with long memory errors," *Journal of Time Series Analysis*, vol. 18, no. 4, pp. 341–354, 1997.

[64] F. Marmol, "Spurious regression theory with nonstationary fractionally integrated processes," *Journal of Econometrics*, vol. 84, no. 2, pp. 233–250, 1998.

[65] W.-J. Tsay and C.-F. Chung, "The spurious regression of fractionally integrated processes," *Journal of Econometrics*, vol. 96, no. 1, pp. 155–182, 2000.

[66] C. R. Nelson and H. Kang, "Spurious periodicity in inappropriately detrended time series," *Econometrica*, vol. 49, no. 3, pp. 741–751, 1981.

[67] D. Dickey and S. Pantula, "Determining the order of differencing in autoregressive processes," *Journal of Business & Economic Statistics*, vol. 5, no. 4, pp. 455–461, 1987.

[68] Y. Sun, "Spurious regressions between stationary generalized long memory processes," *Economics Letters*, vol. 90, no. 3, pp. 446–454, 2006.

[69] T. Mikosch and C. Stărică, "Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects," *The Review of Economics and Statistics*, vol. 86, no. 1, pp. 378–390, 2004.

[70] T. Mikosch and C. de Vries, *Tail Probabilities for Regression Estimators*, Laboratory of Actuarial Mathematics, University of Copenhagen, 2006.

[71] E. Furman and R. Zitikis, "General stein-type decompositions of covariances: revisiting the capital asset pricing model," in *Proceedings of the 58th Annual Meeting of the Midwest Finance Association*, vol. 6, pp. 1–8, Chicago, Ill, USA, 2009.

[72] E. Fama, "The behavior of stock-market prices," *Journal of Business*, vol. 38, no. 1, p. 34, 1965.

[73] U. Hasseler, "Simple regressions with linear time trends," *Journal of Time Series Analysis*, vol. 21, no. 1, pp. 27–32, 2000.

[74] T.-H. Kim, S. Lee, and P. Newbold, "Spurious regressions with processes around linear trends or drifts," 2003, Discussion Papers in Economics.

[75] A. E. Noriega and D. Ventosa-Santaulària, "Spurious regression under broken-trend stationarity," *Journal of Time Series Analysis*, vol. 27, no. 5, pp. 671–684, 2006.

[76] A. E. Noriega and D. Ventosa-Santaulària, "Spurious regression and trending variables," *Oxford Bulletin of Economics & Statistics*, vol. 69, no. 3, pp. 439–444, 2007.

[77] U. Hassler, "Spurious regressions when stationary regressors are included," *Economics Letters*, vol. 50, no. 1, pp. 25–31, 1996.

[78] U. Hassler, "Nonsense regressions due to neglected time-varying means," *Statistical Papers*, vol. 44, no. 2, pp. 169–182, 2003.

[79] W. Newey and K. West, "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, pp. 703–708, 1987.

[80] D. Ventosa-Santaulària, *Spurious Instrumental Variables*, Communications in Statistics. Theory and Methods, 4th edition, 2009.

[81] D. E. A. Giles, "Spurious regressions with time-series data: further asymptotic results," *Communications in Statistics. Theory and Methods*, vol. 36, no. 5–8, pp. 967–979, 2007.

[82] C. M. Jarque and A. K. Bera, "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Economics Letters*, vol. 6, no. 3, pp. 255–259, 1980.

[83] T. S. Breusch and A. R. Pagan, "The Lagrange multiplier test and its applications to model specification in econometrics," *The Review of Economic Studies*, vol. 47, no. 1, pp. 239–253, 1980.

[84] L. G. Godfrey, *Misspecification Tests in Econometrics: The Lagrange Multiplier Principle and Other Methods*, Cambridge University Press, Cambridge, UK, 1988.

[85] D. Ventosa-Santaulària and J. E. Vera-Valdés, "Granger-causality in the presence of structural breaks," *Economics Bulletin*, vol. 3, no. 61, pp. 1–14, 2008.

[86] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.

[87] Y.-S. Lee, T.-H. Kim, and P. Newbold, "Spurious nonlinear regressions in econometrics," *Economics Letters*, vol. 87, no. 3, pp. 301–306, 2005.

[88] J. B. Ramsey, "Tests for specification errors in classical linear least-squares regression analysis," *Journal of the Royal Statistical Society. Series B*, vol. 31, pp. 350–371, 1969.

[89] A. I. McLeod and W. K. Li, "Diagnostic checking ARMA time series models using squared-residual autocorrelations," *Journal of Time Series Analysis*, vol. 4, no. 4, pp. 269–273, 1983.

[90] D. M. Keenan, "A Tukey nonadditivity-type test for time series nonlinearity," *Biometrika*, vol. 72, no. 1, pp. 39–44, 1985.

[91] H. White, "An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks," in *Proceedings of the International Joint Conference on Neural Networks*, vol. 2, pp. 451–455, IEEE Press, Washington, DC, USA, 1989.

[92] H. White, *Estimation, Inference and Specification Analysis*, Cambridge University Press, New York, NY, USA, 1992.

[93] J. D. Hamilton, "A parametric approach to flexible nonlinear inference," *Econometrica*, vol. 69, no. 3, pp. 537–573, 2001.

[94] A. Noriega and D. Ventosa-Santaulària, "Spurious cointegration: the Engle-Granger test in the presence of structural breaks," Working paper no. 2006-12, Banco de México, 2006.

[95] R. Valkanov, "Long-horizon regressions: theoretical results and applications," *Journal of Financial Economics*, vol. 68, no. 2, pp. 201–232, 2003.

[96] J. Lee, "Fractionally integrated long horizon regressions," *Studies in Nonlinear Dynamics & Econometrics*, vol. 11, no. 1, 2007.

[97] S. J. Leybourne and P. Newbold, "Spurious rejections by cointegration tests induced by structural breaks," *Applied Economics*, vol. 35, no. 9, pp. 1117–1121, 2003.

[98] P. C. B. Phillips and B. E. Hansen, "Statistical inference in instrumental variables regression with I (1) processes," *Review of Economic Studies*, vol. 57, no. 1, pp. 99–125, 1990.

[99] B. Hansen and P. Phillips, "Estimation and inference in models of cointegration: a simulation study," *Advances in Econometrics*, vol. 8, pp. 225–248, 1990.

[100] P. M. Robinson, "Efficient tests of nonstationary hypotheses," *Journal of the American Statistical Association*, vol. 89, no. 428, pp. 1420–1437, 1994.

[101] D. A. Dickey and W. A. Fuller, "Likelihood ratio statistics for autoregressive time series with a unit root," *Econometrica*, vol. 49, no. 4, pp. 1057–1072, 1981.

[102] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root?" *Journal of Econometrics*, vol. 54, no. 1–3, pp. 159–178, 1992.

[103] G. Elliott, T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, vol. 64, no. 4, pp. 813–836, 1996.

[104] P. C. B. Phillips and P. Perron, "Testing for a unit root in time series regression," *Biometrika*, vol. 75, no. 2, pp. 335–346, 1988.

[105] S. Ng and P. Perron, "Lag length selection and the construction of unit root tests with good size and power," *Econometrica*, vol. 69, no. 6, pp. 1519–1554, 2001.

[106] P. Perron and S. Ng, "Useful modifications to some unit root tests with dependent errors and their local asymptotic properties," *The Review of Economic Studies*, vol. 63, no. 3, pp. 435–463, 1996.

[107] P. Perron, "Testing for a unit root in a time series with a changing mean," *Journal of Business & Economic Statistics*, pp. 153–162.

[108] A. Montañés and M. Reyes, "Effect of a shift in the trend function on Dickey-Fuller unit root tests," *Econometric Theory*, vol. 14, no. 3, pp. 355–363, 1998.

[109] A. Montañés and M. Reyes, "The asymptotic behaviour of the Dickey-Fuller tests under the crash hypothesis," *Statistics & Probability Letters*, vol. 42, no. 1, pp. 81–89, 1999.

[110] S. J. Leybourne and P. Newbold, "Behavior of Dickey-Fuller t-tests when there is a break under the alternative hypothesis," *Econometric Theory*, vol. 16, no. 5, pp. 779–789, 2000.

[111] A. Sen, "Behaviour of Dickey-Fuller F-tests under the trend-break stationary alternative," *Statistics & Probability Letters*, vol. 55, no. 3, pp. 257–268, 2001.

[112] A. Sen, "On unit-root tests when the alternative is a trend-break stationary process," *Journal of Business & Economic Statistics*, vol. 21, no. 1, pp. 174–184, 2003.

[113] T.-H. Kim, S. Leybourne, and Pa. Newbold, "Behaviour of Dickey-Fuller unit-root tests under trend misspecification," *Journal of Time Series Analysis*, vol. 25, no. 5, pp. 755–764, 2004.

[114] S. Hamori and A. Tokihisa, "Testing for a unit root in the presence of a variance shift," *Economics Letters*, vol. 57, no. 3, pp. 245–253, 1997.

[115] S. J. Leybourne, T. C. Mills, and P. Newbold, "Spurious rejections by Dickey-Fuller tests in the presence of a break under the null," *Journal of Econometrics*, vol. 87, no. 1, pp. 191–203, 1998.

[116] T.-H. Kim, S. Leybourne, and P. Newbold, "Unit root tests with a break in innovation variance," *Journal of Econometrics*, vol. 109, no. 2, pp. 365–387, 2002.

[117] A. Sen, "Behaviour of Dickey-Fuller tests when there is a break under the unit root null hypothesis," *Statistics and Probability Letters*, vol. 78, no. 6, pp. 622–628, 2008.

[118] E. Zivot and D. W. K. Andrews, "Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis," *Journal of Business and Economic Statistics*, vol. 10, pp. 251–270, 1992.

[119] J. L. Carrion-i-Silvestre and A. Sansó, "Joint hypothesis specification for unit root tests with a structural break," *The Econometrics Journal*, vol. 9, no. 2, pp. 196–224, 2006.

[120] J. Bai and P. Perron, "Estimating and testing linear models with multiple structural changes," *Econometrica*, vol. 66, no. 1, pp. 47–78, 1998.

[121] B. B. Mandelbrot and M. S. Taqqu, "Robust R/S analysis of long-run serial correlation," in *Proceedings of the 42nd Session of the International Statistical Institute (ISI '79)*, vol. 48, pp. 69–99, 1979.

[122] A. Lo, "Long-term memory in stock market prices," *Econometrica*, pp. 1279–1313, 1991.

[123] T. Mikosch and C. Stărică, "Is it really long memory we see in financial returns," *Extremes and Integrated Risk Management*, vol. 12, pp. 149–168, 2000.

[124] V. Teverovsky and M. Taqqu, "Testing for long-range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator," *Journal of Time Series Analysis*, vol. 18, no. 3, pp. 279–304, 1997.

[125] Y.- Liu, M. Pan, and L. Hsueh, "A modified R/S analysis of long-term dependence in currency futures prices," *Journal of International Financial Markets, Institutions & Money*, vol. 3, no. 2, pp. 97–113, 1993.

[126] I. Lobato and P. M. Robinson, "Averaged periodogram estimation of long memory," *Journal of Econometrics*, vol. 73, no. 1, pp. 303–324, 1996.

[127] L. Giraitis, P. Kokoszka, and R. Leipus, "Testing for long memory in the presence of a general trend," *Journal of Applied Probability*, vol. 38, no. 4, pp. 1033–1054, 2001.

[128] L. Giraitis, P. Kokoszka, R. Leipus, and G. Teyssière, "Rescaled variance and related tests for long memory in volatility and levels," *Journal of Econometrics*, vol. 112, no. 2, pp. 265–294, 2003.

[129] I. Berkes, L. Horváth, P. Kokoszka, and Q.-M. Shao, "On discriminating between long-range dependence and changes in mean," *The Annals of Statistics*, vol. 34, no. 3, pp. 1140–1165, 2006.

[130] A. Zhang, R. Gabrys, and P. Kokoszka, "Discriminating between long memory and volatility shifts," *Statistics*, vol. 36, no. 4, pp. 253–275, 2007.

[131] A. Aue, L. Horváth, M. Hušková, and P. Kokoszka, "Testing for changes in polynomial regression," *Bernoulli*, vol. 14, no. 3, pp. 637–660, 2008.

[132] A. Jach and P. Kokoszka, "Wavelet-domain test for long-range dependence in the presence of a trend," *Statistics*, vol. 42, no. 2, pp. 101–113, 2008.

[133] J. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, NJ, USA, 1994.

[134] R. King, C. Plosser, J. Stock, and M. Watson, "Stochastic trends and economic fluctuations," *The American Economic Review*, vol. 81, no. 4, pp. 819–840, 1991.

[135] J. Cochrane, "Permanent and transitory components of GNP and stock prices," *The Quarterly Journal of Economics*, vol. 109, no. 1, pp. 241–265, 1994.

[136] A. Bernard and S. Durlauf, "Convergence in International output," *Journal of Applied Econometrics*, vol. 10, no. 2, pp. 97–108, 1995.

[137] A. B. Bernard and S. N. Durlauf, "Interpreting tests of the convergence hypothesis," *Journal of Econometrics*, vol. 71, no. 1-2, pp. 161–173, 1996.

[138] S. Durlauf, "On the convergence and divergence of growth rates," *The Economic Journal*, vol. 106, no. 437, pp. 1016–1018, 1996.

[139] Z. Koustas and W. Veloce, "Unemployment hysteresis in Canada: an approach based on long-memory time series models," *Applied Economics*, vol. 28, no. 7, pp. 823–831, 1996.

[140] S. Durlauf and D. Quah, *The New Empirics of Economic Growth*, National Bureau of Economic Research Cambridge, Cambridge, Mass, USA, 1998.

[141] Q. Li and D. Papell, "Convergence of international output time series evidence for 16 OECD countries," *International Review of Economics and Finance*, vol. 8, no. 3, pp. 267–280, 1999.

[142] S. N. Durlauf, "Manifesto for a growth econometrics," *Journal of Econometrics*, vol. 100, no. 1, pp. 65–69, 2001.

[143] P. Aghion and S. N. Durlauf, *Handbook of Economic Growth, Volume 1A*, North-Holland, Amsterdam, The Netherlands, 2005.

[144] J. Campbell and P. Perron, "What macroeconomist should know about unit roots," in *NBER Macroeconomics Annual*, O. Blanchard and S. Fisher, Eds., pp. 141–201, MIT Press, Cambridge, Mass, USA, 1991.

[145] J. Faust and E. M. Leeper, "When do long-run identifying restrictions give reliable results?" *Journal of Business & Economic Statistics*, vol. 15, no. 3, pp. 345–353, 1997.

[146] P. J. Coe and J. M. Nason, "Long-run monetary neutrality and long-horizon regressions," *Journal of Applied Econometrics*, vol. 19, no. 3, pp. 355–373, 2004.

[147] J. Taylor, "Estimation and control of a macroeconomic model with rational expectations," *Econometrica*, vol. 47, no. 5, pp. 1267–1286, 1979.

[148] D. Corbae and S. Ouliaris, "Cointegration and tests of purchasing power parity," *The Review of Economics and Statistics*, vol. 70, no. 3, pp. 508–511, 1988.

[149] B. Diba and H. Grossman, "Explosive rational bubbles in stock prices?" *The American Economic Review*, vol. 78, no. 3, pp. 520–530, 1988.

[150] Y. Cheung, "Long memory in foreign-exchange rates," *Journal of Business & Economic Statistics*, pp. 93–101, 1993.

[151] N. Crato and P. J. F. de Lima, "Long-range dependence in the conditional variance of stock returns," *Economics Letters*, vol. 45, no. 3, pp. 281–285, 1994.

[152] W. Enders and S. Hurn, "Theory and tests of generalized purchasing-power parity: common trends and real exchange rates in the Pacific Rim," *Review of International Economics*, vol. 2, no. 2, pp. 179–190, 1994.

[153] R. Garcia and P. Perron, "An analysis of the real interest rate under regime shifts," *The Review of Economics and Statistics*, vol. 78, no. 1, pp. 111–124, 1996.

[154] S. E. Culver and D. H. Papell, "Is there a unit root in the inflation rate? Evidence from sequential break and panel data models," *Journal of Applied Econometrics*, vol. 12, no. 4, pp. 435–444, 1997.

[155] F. J. Breidt, N. Crato, and P. de Lima, "The detection and estimation of long memory in stochastic volatility," *Journal of Econometrics*, vol. 83, no. 1-2, pp. 325–348, 1998.

[156] I. N. Lobato and N. E. Savin, "Real and spurious long-memory properties of stock-market data," *Journal of Business & Economic Statistics*, vol. 16, no. 3, pp. 261–283, 1998.

[157] Z. Zhu, "The random walk of stock prices: evidence from a panel of G-7 countries," *Applied Economics Letters*, vol. 5, no. 7, pp. 411–413, 1998.

[158] R. C. Cati, M. G. P. Garcia, and P. Perron, "Unit roots in the presence of abrupt governmental interventions with an application to Brazilian data," *Journal of Applied Econometrics*, vol. 14, no. 1, pp. 27–56, 1999.

[159] J. T. Cuddington and H. Liang, "Purchasing power parity over two centuries?" *Journal of International Money and Finance*, vol. 19, no. 5, pp. 753–757, 2000.

[160] T. Mikosch, *Modeling Dependence and Tails of Financial Time Series*, Extreme Values in Finance, Telecommunications, and the Environment, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2003.

[161] P. Pedroni, "Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis," *Econometric Theory*, vol. 20, no. 3, pp. 597–625, 2004.

[162] A. M. R. Taylor, "Fluctuation tests for a change in persistence," *Oxford Bulletin of Economics & Statistics*, vol. 67, no. 2, pp. 207–230, 2005.

[163] G. Cavaliere and A. M. R. Taylor, "Testing for a change in persistence in the presence of a volatility shift," *Oxford Bulletin of Economics & Statistics*, vol. 68, supplement 1, pp. 761–781, 2006.

[164] E. Furman and R. Zitikis, "Weighted risk capital allocations," *Insurance: Mathematics & Economics*, vol. 43, no. 2, pp. 263–269, 2008.

[165] E. Furman and R. Zitikis, "Weighted pricing functionals," in *Proceedings of the 43rd Actuarial Research Conference (ARC '09)*, pp. 1–31, University of Regina, Saskatchewan, Canada, 2009.

[166] G. S. Alogoskoufis and R. Smith, "The Phillips curve, the persistence of inflation, and the Lucas critique: evidence from exchange rate regimes," *The American Economic Review*, vol. 81, no. 5, pp. 1254–1275, 1991.

[167] M. D. D. Evans and K. K. Lewis, "Do expected shifts in inflation affect estimates of the long-run Fisher relation?" *The Journal of Finance*, vol. 50, no. 1, pp. 225–253, 1995.

[168] J. Fuhrer and G. Moore, "Inflation persistence," *The Quarterly Journal of Economics*, vol. 110, no. 1, pp. 127–159, 1995.

[169] W. J. Crowder, "The long-run Fisher relation in Canada," *The Canadian Journal of Economics*, vol. 30, no. 4b, pp. 1124–1142, 1997.

[170] S. B. DeLoach, "More evidence in favor of the Balassa-Samuelson hypothesis," *Review of International Economics*, vol. 9, no. 2, pp. 336–342, 2001.

[171] I. Drine and C. Rault, "Do panel data permit the rescue of the Balassa-Samuelson hypothesis for Latin American countries?" *Applied Economics*, vol. 35, no. 3, pp. 351–359, 2003.

[172] J. Ricardo Faria and M. León-Ledesma, "Testing the Balassa-Samuelson effect: implications for growth and the PPP," *Journal of Macroeconomics*, vol. 25, no. 2, pp. 241–253, 2003.

[173] G. Bårdsen, E. S. Jansen, and R. Nymoen, "Econometric evaluation of the New Keynesian Phillips curve," *Oxford Bulletin of Economics & Statistics*, vol. 66, supplement, pp. 671–686, 2004.

[174] Y. Sun and P. C.B. Phillips, "Understanding the fisher equation," *Journal of Applied Econometrics*, vol. 19, no. 7, pp. 869–886, 2004.

[175] J. Galí, M. Gertler, and J. López-Salido, *Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve*, Banco de España, 2005.

[176] P. C. B. Phillips, "Econometric analysis of Fisher's equation," *American Journal of Economics and Sociology*, vol. 64, no. 1, pp. 125–168, 2005.