Co-summability

From Linear to Non-linear Co-integration[*]

Vanessa Berenguer-Rico
University of Oxford and Institute for New Economic Thinking
Jesus Gonzalo
Universidad Carlos III de Madrid

October 9, 2014

Abstract

Co-integration plays a fundamental role in the econometric analysis of linear relationships among persistent economic time series. Nonetheless, nonlinearities are often encountered in modern macroeconometric models that account for more flexible relationships. In a nonlinear world, however, the concepts order of integration and co-integration are not readily applicable. The inherent linearity in the order of integration idea invalidates its use to characterize nonlinear persistent and/or nonstationary processes and this, in turn, implies that co-integration cannot be directly extended to study nonlinear relationships.

To overcome these hindrances, Berenguer-Rico and Gonzalo (2014) formalized the concept order of summability of a stochastic process, which generalizes the order of integration idea to nonlinear time series. In this paper, the order of summability is used to extend co-integration to non-linear models. Specifically, we formalise the idea of co-summability and propose a residual-based statistic to test for it. The statistic can also be seen as a misspecification testing procedure and is based on the order of summability of the error term. The performance of the test is studied via Monte Carlo experiments. Finally, the practical strength of co-summability theory is shown through two empirical applications. In particular, asymmetric preferences of central bankers and the environmental Kuznets curve hypothesis are studied through the lens of co-summability.

Key words: Balancedness, Co-integration, Co-summability, Non-linear Co-integration, Misspecification, Non-linear Processes, Non-stationarity, Order of Summability, Persistence

JEL Codes: C01, C22

[*]The authors acknowledge comments from J.Ll. Carrion-i-Silvestre, J.J. Dolado, B. Nielsen, G. Perez-Quiros, P.C.B. Phillips, J. Pitarakis, C. Velasco, and seminar and conference participants at Nuffield/INET, NES/CEFIR, ECARES, UCM, XXXVI SAE, 5th CFE, 2012 Tsinghua Econometrics, SETA 2012, Southampton Spring Econometrics Event and ASSA Meetings 2013 for their helpful insights. Financial support from the University of Oxford, Institute for New Economic Thinking, Spanish Ministerio de Ciencia e Innovación (grants ECO2010-19357 and Consolider-2010), Comunidad de Madrid (grant Excelecon) and Bank of Spain (grant ER program) is gratefully acknowledged.

1 Introduction

Co-integration theory has received a great deal of attention from economists and econometricians. Through the idea of cointegration, linear equilibrium relationships between macroeconomic variables hypothesized by economic theorists were statistically supported from a time series analysis perspective. More recently, researchers have ventured into the non-linear world to provide richer descriptions of economic phenomena. However, the ideas of integration and co-integration cannot be directly used to analyse non-linear equilibrium relationships among persistent variables as these concepts do not properly apply. To be more precise, the order of integration of nonlinear transformations of persistent processes may not be well defined. This failure of applicability of the definition of order of integration implies that the concept of co-integration cannot be directly extended to non-linear long run relationships. As already pointed out by Granger (1995), this originates a clear need for theoretically valid and empirically useful concepts that generalise those of integration and co-integration.

This paper proposes to use the idea of order of summability formalised by Berenguer-Rico and Gonzalo (2014). It was conceived to deal both theoretically and empirically with non-linear transformations of heterogeneous and/or persistent processes. By making use of this new concept, co-integration theory can be generalised through the idea of co-summability to non-linear relationships. By taking advantage of the order of summability estimator, co-summability can be empirically studied. To infer if a postulated relationship is co-summable, this paper proposes a residual based test, therefore, an estimate of the errors is needed. Parametric and non-parametric approaches to estimate non-linear long run relationships are available in the literature. Park and Phillips (1999, 2001) and Wang and Phillips (2009) develop parametric and non-parametric methods, respectively, from an integrated processes perspective. Alternatively, Karlsen, Myklebust and Tjϕstheim (2007) and Schienle (2011) analyse non-parametric estimation in a recurrent Markov chains setup. Notwithstanding, all these studies assume that the regression model specifies a co-integrating relation, something that should be tested in practice. There have been some –rather limited– proposals in this direction –see, for example, Kasparis (2008) or Choi and Saikkonen (2010). The test proposed in this paper contributes to this literature by giving more insights into the degree of persistence and heterogeneity of the residuals (hence into the degree of misspecification), while at the same time delivering reasonable properties in terms of size an power compared to the existing tests.

To show the empirical strength of co-summability theory, the proposed test is put into practice with two different empirical applications where non-linear transformations of persistent processes occur. Specifically, asymmetric preferences of central bankers and the environmental Kuznets curve are analysed. The former hypothesis is translated in the literature into non-linear Taylor rules when conducting monetary policy –see, for instance, Clarida and Gertler (1997) or Dolado, María-Dolores and Naveira (2005). These non-linearities and the fact that the variables involved in this type of rules are found to be persistent make co-summability appropriate in this context. The latter hypothesis, the environmental Kuznets curve, postulates an inverted U-shaped relationship between pollution and economic development, usually measured by CO_2 emissions and GDP, respectively –see Dasgupta et al. (2001) or Brock and Taylor (2005) for an overview. Again, this non-linear relationship, typically approximated by a polynomial function, jointly with the well documented persistence of these two measures, make this hypothesis another natural economic context where co-summability theory rightly fits. The empirical findings provide new insights for the econometric treatment of these two hypotheses. In the Taylor rule case, the linear specification does not define a long run relationship –co-summability does not hold– thus suggesting a possible misspecification. Following the asymmetric preferences of central bankers literature, we find that a threshold Taylor rule is not rejected –co-summability holds. Specifically, it is found that the Federal Reserve reacts very asymmetrically to recessions and expansions. With respect to the environmental Kuznets curve, favourable evidence is found when variables are included in logarithms and the polynomial function is of third degree.

The paper is organised as follows. In Section 2, co-summability is formally defined and discussed through two economic examples. Section 3 describes the model to be considered and Section 4 proposes a test for strong co-summability. The performance of the test is studied via simulations. In Section 5, two empirical exercises are carried out: asymmetric preferences of central bankers and the environmental Kuznets curve hypothesis are analyzed. Section 6 finishes with some concluding remarks. The proofs are collected in the Appendix.

2 Co-summability

The subsequent theory relies on the idea of order of summability of stochastic processes. It was first introduced in a heuristic way by Gonzalo and Pitarakis (2006) and subsequently formalised in Berenguer-Rico and Gonzalo (2014) –BG hereafter. A slightly simplified version of the definition in BG is used in this paper[1].

Definition 1 *: A stochastic process* $\{y_t : t \in \mathbb{N}\}$ *is said to be summable of order* δ*, or* $S(\delta)$*, if*

$$S_n = \frac{1}{n^{\frac{1}{2}+\delta}} \sum_{t=1}^{n} y_t = O_p(1),$$

but not $o_p(1)$ *as* $n \to \infty$*.*

[1] In BG, S_n is defined as

$$S_n = \frac{1}{n^{\frac{1}{2}+\delta}} L(n) \sum_{t=1}^{n} (y_t - m_t),$$

where $L(n)$ is a slowly varying function and m_t a deterministic sequence.

The order of summability, δ, gives a summary measure of the stochastic properties –persistence and evolution of the variance– of y_t without relying on a particular data generating process. Hence, contrary to the order of integration, it is able to characterize nonlinear processes. It generalizes the idea of order of integration in the sense that any $I(d)$ process, $d \geq 0$, is $S(d)$ –see BG for a formal argument on the relationship between $I(d)$ and $S(d)$. In contrast to the order of integration, the order of summability can also easily characterize nonlinear processes. As a typical example, consider the square of a random walk

$$y_t^2 = y_{t-1}^2 + y_{t-1}\varepsilon_t + \varepsilon_t^2,$$

where $\varepsilon_t \sim i.i.d.(0,1)$ and $y_0 = 0$. Findining the number d that makes $\Delta^d y_t^2$ to be an $I(0)$ process poses a number of conceptual difficulties. If, instead, the order of summability is used and given that

$$\frac{1}{n^2}\sum_{t=1}^{n} y_t^2 \rightsquigarrow \int_0^1 W(r)^2\, dr,$$

we have $y_t^2 \sim S(1.5)$. Therefore, the order of summability of y_t^2 can be easily characterized through the asymptotic behavior of its sum. Other examples, such as cross-products, threshold, logistic or logarithmic transformations are analyzed in BG; Table 1 summarises them.

Once a generalization of the order of integration for non-linear processes is available, an extension of co-integration to nonlinear relationships can be made through the concept of co-summability.

Definition 2 *: Two summable stochastic processes, $y_t \sim S(\delta_y)$ and $x_t \sim S(\delta_x)$, are said to be co-summable if there exists $f(x_t) \sim S(\delta_y)$ such that $u_t = y_t - f(x_t)$ is $S(\delta_u)$, with $\delta_u = \delta_y - \delta$ and $\delta > 0$. In short, $(y_t, x_t) \sim CS(\delta_y, \delta)$.*

Table 1: Examples: $I(d)$ vs $S(\delta)$

DGP	$I(d)$	$S(\delta)$	*DGP*	$I(d)$	$S(\delta)$
$y_{1t} \sim i.i.d.F \in D(\alpha)$	$I(?)$	$S((2-\alpha)/2\alpha)$	$y_{8t} = 1(v_t \leq \gamma)\pi_t$	$I(?)$	$S(1)$
$y_{2t} = z + \varepsilon_t$	$I(?)$	$S(1/2)$	$y_{9t} = e^{-\pi_t^2}$	$I(?)$	$S(1/2)$
$y_{3t} \sim I(d)$	$I(d)$	$S(d)$	$y_{10t} = 1/(1+\pi_t^2)$	$I(?)$	$S(1/2)$
$y_{4t} = \pi_t\eta_t$	$I(?)$	$S(1/2)$	$y_{11t} = \log(\lvert\pi_t\rvert)$	$I(?)$	$S(1/2)$
$y_{5t} = \pi_t\eta_t^2$	$I(?)$	$S(1)$	$y_{12t} = (1+e^{-\pi_t})^{-1}$	$I(?)$	$S(1/2)$
$y_{6t} = \pi_t^2$	$I(?)$	$S(3/2)$	$y_{13t} = \rho_t y_{12,t-1} + \varepsilon_t$	$I(?)$	$S(\delta(n))$
$y_{7t} = \pi_t^\lambda,\ \lambda = 1,2...$	$I(?)$	$S((1+\lambda)/2)$	$y_{14t} = \phi y_{13,t-1} + \varepsilon_t;\quad \phi > 1$	$I(?)$	$S(\delta(n))$

$D(\alpha)$ denotes the domain of attraction of an α-stable law with $\alpha \in (0,2]$; $z \sim N(0,1)$; $\varepsilon_t \sim i.i.d.(0,1)$; $\pi_t = \pi_{t-1} + \varepsilon_t$ and $\pi_0 = 0$; $\eta_t \sim i.i.d.(0,1)$; $v_t \sim i.i.d.(0,1)$; $\rho_t \sim i.i.d.(1,1)$. z, ε_t, η_t, v_t, and ρ_t are independent of each other. In this Table, I(d) is understood as the definition given by Engle and Granger (1987), p.252.

A couple of economic examples, to be analyzed empirically below, will help in motivating the theoretical and practical need for an extension of the idea of cointegration to nonlinear models and show that co-summability is able to cover this need.

Example 1 *: Central Bankers with Asymmetric Preferences*

Consider a central bank with asymmetric preferences with respect to deviations of inflation or output from some particular target level. Different modelisations of this hypothesis based on Taylor rules can be found in the literature. For instance, Clarida and Gertler (1997) study the following threshold type of Taylor rule for the Bundesbank

$$i_t = \theta_0 + \theta_1 \tilde{\pi}_t 1(\tilde{\pi}_t > 0) + \theta_2 \tilde{\pi}_t 1(\tilde{\pi}_t \leq 0) + \theta_3 \tilde{y}_t 1(\tilde{\pi}_t > 0) + \theta_4 \tilde{y}_t 1(\tilde{\pi}_t \leq 0) + u_t, \quad (1)$$

where i_t denotes interest rates, $\tilde{\pi}_t$ are deviations from the inflation target, and $\tilde{y}_t$ is the output gap. On the other hand, Dolado, María-Dolores and Naveira (2005), allowing for a non-linear Phillips curve, derive the following type of optimal monetary policy rule

$$i_t = \theta_0 + \theta_1 \tilde{\pi}_t + \theta_2 \tilde{y}_t + \theta_3 \tilde{\pi}_t \tilde{y}_t + u_t. \quad (2)$$

In both cases, extending co-integration will be troublesome using the $I(d)$ framework. Even if it can be said that i_t, $\tilde{\pi}_t$, and $\tilde{y}_t$ are $I(d_i)$, $I(d_{\tilde{\pi}})$, and $I(d_{\tilde{y}})$, respectively, the order of integration of $\tilde{\pi}_t 1(\tilde{\pi}_t \geq 0)$ or $\tilde{\pi}_t \tilde{y}_t$ may not be well defined. Nevertheless, the generality of the order of summability makes it suitable to be used in both situations, and hence, co-summability would deliver a well defined long run relationship.

Example 2 *: Environmental Kuznets Curve*

The environmental Kuznets curve indicates an inverted-U relationship between pollution and economic development –see Dasgupta et al. (2001) or Brock and Taylor (2005) for an overview. The usual shape given to this relationship is of a polynomial type. Consider the simplest

$$p_t = \theta_0 + \theta_1 y_t + \theta_2 y_t^2 + u_t,$$

where p_t is a measure of pollution and y_t is a measure of income, typically CO_2 and GDP, respectively. Using the standard notion of co-integration in this model will be again problematic. Even if it is known that y_t is $I(d_y)$, the order of integration of y_t^2 could not be well defined. As it has been emphasised herein, the order of summability can help to overcome this pitfall and therefore co-summability provides a framework where properly defining this long run relation.

3 The Model

The co-summable relationship to be analysed in this paper is the one described by the following model, linear in parameters but possibly non-linear in variables,

$$y_t = \theta_0 f(x_t) + u_t, \tag{3}$$

where $f : \mathbb{R} \to \mathbb{R}$, x_t and y_t are known by the researcher. Relationship (3) can be considered to be an approximation to a more general relationship $y_t = g(x_t, \theta_0) + u_t$, which will be better than the standard approximation considered in co-integration theory –linear in parameters and variables. Indeed, if g is such that a Taylor expansion applies, then only considering a linear relation could unbalance the model if, for instance, a higher order polynomial were a better approximation. Moreover, model (3) is empirically very rich. To facilitate the exposition, the bivariate case (y_t, x_t) will be considered but the extension to a multivariate x_t or to additively separable multiple regression models can be easily adapted.

Consider that the following least squares regression is carried out

$$y_t = \hat{\theta}_n f(x_t) + \hat{u}_t, \quad t = 1, ..., n. \tag{4}$$

Assumption 1

(a) For some $\delta_f > 0$ and $\delta_u \geq 0$,

$$\left(\frac{1}{n^{1/2+\delta_f}} \sum_{t=1}^{[nr]} f(x_t), \frac{1}{n^{1/2+\delta_u}} \sum_{t=1}^{[nr]} u_t \right) \implies (D_f(r), D_u(r)),$$

where $(D_f(r), D_u(r))$ is a vector of stochastic processes defined on $r \in [0, 1]$.

(b)

$$n^{\delta_f - \delta_u} \left(\hat{\theta}_n - \theta_0 \right) \implies D_\theta,$$

where D_θ is some random variable.

Assumption 1 is a general condition that will allow us to work under both the null hypothesis of strong co-summability, SC, $H_o : \delta_u = 0$, and the alternative of no strong co-summability, NSC, $H_a : \delta_u > 0$. This hypotheses framework, along the lines of the KPSS test for co-integration in the $I(d)$ world, allows us to construct a functional form misspecification test. Under the null hypothesis, $H_o : \delta_u = 0$, Assumption 1 implies that

$$\frac{1}{n^{1/2}} \sum_{t=1}^{[nr]} u_t \implies D_u(r) \quad \text{and} \quad n^{\delta_f} \left(\hat{\theta}_n - \theta_0 \right) \implies D_\theta.$$

These conditions will be satisfied under different data generating processes describing u_t and $f(x_t)$.

Several simple examples in which these assumptions are met are put forward next. The assumptions in these examples, referring to the error term u_t and $f(x_t)$, can be made more general as several authors have shown. For expository and explanatory convenience, we describe these simple cases next.

3.1 Examples under the null hypothesis

Example 3 *:* Linear Cointegration

(a) $y_t = \theta_0 x_t + u_t$.

(b) Let $\mathcal{F}_{nt}$, $t = 0, ..., n$, such that $(u_t, \mathcal{F}_{nt})$ is a martingale difference sequence with $E\left(u_t^2 | \mathcal{F}_{n,t-1}\right) = \sigma^2$ $a.s.$ for all $t = 1, ..., n$ and $\sup_{1 \le t \le n} \left(|u_t|^q | \mathcal{F}_{n,t-1}\right) < \infty$ $a.s.$ for some $q > 2$.

(c)

$$\left(\frac{x_{[nr]}}{n^{1/2}}, \frac{1}{n^{1/2}} \sum_{t=1}^{[nr]} u_t \right) \implies (W(r), V(r)),$$

where $(W(r), V(r))$ is a vector Brownian motion.

It can be easily seen that conditions in Assumptions 1 are satisfied in this case. In particular, $\delta_f = 1$ and $\delta_u = 0$, so that,

$$\left(\frac{1}{n^{3/2}} \sum_{t=1}^{[nr]} x_t, \frac{1}{n^{1/2}} \sum_{t=1}^{[nr]} u_t \right) \implies \left(\int_0^r W(r)\, dr, V(r) \right).$$

and

$$n\left(\hat{\theta}_n - \theta_0\right) \implies \frac{\int_0^1 W(r)\, dV(r)}{\int_0^1 W^2(r)\, dr}.$$

Example 4 *:* Park and Phillips (1999)

(a) $y_t = \theta_0 f(x_t) + u_t$.

(b) Let $\mathcal{F}_{nt}$, $t = 0, ..., n$, such that $(u_t, \mathcal{F}_{nt})$ is a martingale difference sequence with $E\left(u_t^2 | \mathcal{F}_{n,t-1}\right) = \sigma^2$ $a.s.$ for all $t = 1, ..., n$ and $\sup_{1 \le t \le n} \left(|u_t|^q | \mathcal{F}_{n,t-1}\right) < \infty$ $a.s.$ for some $q > 2$.

(b) $x_t = x_{t-1} + w_t$ with $w_t = \psi(L) \varepsilon_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k}$, $\varepsilon_t \sim i.i.d.\left(0, \sigma_\varepsilon^2\right)$, $\psi(1) \neq 0$, $x_0 = O_p(1)$, $\sigma_\varepsilon^2 < \infty$ and $\sum_{k=0}^{\infty} k^{1/2} |\psi_k| < \infty$.

(c) u_t is independent of w_t.

(d) $f(.)$ is asymptotically homogeneous with asymptotic order $v(\sqrt{n}) = n^{\delta_f - 1/2}$ and limit homogeneous function $h(.)$.

Conditions in Assumptions 1 are satisfied in this example too. In particular, from Park and Phillips (1999),

$$\left(\frac{1}{n^{1/2+\delta_f}} \sum_{t=1}^{[nr]} f(x_t), \frac{1}{n^{1/2}} \sum_{t=1}^{[nr]} u_t \right) \implies \left(\int_0^r h(W(r))\, dr, V(r) \right),$$

and

$$n^{\delta_f}\left(\hat{\theta}_n-\theta_0\right) \Longrightarrow \frac{\int_0^1 f\left(W\left(r\right)\right)dV\left(r\right)}{\int_0^1 f^2\left(W\left(r\right)\right)dr}.$$

3.2 Examples under the alternative

Assumption 1 also considers models under the alternative hypothesis, $H_a : \delta_u > 0$. The following examples describe two concrete alternatives.

Example 5 *:* Spurious Regressions (Phillips, 1986)

(a) $y_t = y_{t-1} + e_{yt}$ with $e_{yt} \sim i.i.d.\left(0,1\right)$ and $y_0 = 0$.

(b) $x_t = x_{t-1} + e_{xt}$ with $e_{xt} \sim i.i.d.\left(0,1\right)$ and $x_0 = 0$.

(c) e_{yt} and e_{xt} are independent of each other.

Under these conditions

$$\left(\frac{1}{n^{3/2}}\sum_{t=1}^{[nr]} x_t, \frac{1}{n^{3/2}}\sum_{t=1}^{[nr]} y_t\right) \Longrightarrow \left(\int_0^r W_x\left(s\right)ds, \int_0^r W_y\left(s\right)ds\right),$$

and

$$\left(\hat{\theta}_n-\theta_0\right) \Longrightarrow \frac{\int_0^1 W_x\left(r\right)W_y\left(r\right)dr}{\int_0^1 W_x^2\left(r\right)dr}.$$

Hence, conditions in Assumption 1 hold with $\delta_f = \delta_u = \delta_y = 1$.

Example 6 *:* Functional Form Misspecification

(a) $y_t = \theta_0 x_t + u_t$

(b) $x_t = x_{t-1} + v_t$

(c) $u_t = \gamma_0 x_t^2 + \varepsilon_t$

(d) $v_t \sim i.i.d.\left(0,1\right)$ and $\varepsilon_t \sim i.i.d.\left(0,1\right)$ are independent of each other.

In this case,

$$\left(\frac{1}{n^{3/2}}\sum_{t=1}^{[nr]} x_t, \frac{1}{n^2}\sum_{t=1}^{[nr]} u_t\right) \Longrightarrow \left(\int_0^r W_x\left(s\right)ds, \gamma_0\int_0^r W_x^2\left(r\right)dr\right),$$

and

$$n^{-1/2}\left(\hat{\theta}_n-\theta_0\right) \Longrightarrow \gamma_0\frac{\int_0^1 W_x^3\left(r\right)dr}{\int_0^1 W_x^2\left(r\right)dr}.$$

Hence, conditions in Assumption 1 hold with $\delta_f = 1$ and $\delta_u = 1.5$.

4 Testing for Strong Co-summability

In this section, we propose a residual-based test for strong co-summability. Our proposal is based on the order of summability of the residuals; given their triangular array nature, in this section, we modify accordingly the definition of order of summability to accomodate them.

Definition 3 *: A triangular array $\{y_{nt} : n \in \mathbb{N}, 1 \leq t \leq n\}$ is said to be summable of order δ, or $S(\delta)$, if*

$$S_n = \frac{1}{n^{\frac{1}{2}+\delta}} \sum_{t=1}^{n} y_{nt}, \tag{5}$$

is $O_p(1)$ but not $o_p(1)$ as $n \to \infty$.

Proposition 1 *: Let $\hat{u}_t$ be the OLS residuals of $y_t = \theta_0 f(x_t) + u_t$.*

(a) Under Assumption 1, if $\delta_u = 0$,

$$\frac{1}{n^{1/2}} \sum_{t=1}^{n} \hat{u}_t = O_p(1).$$

(b) Under Assumption 1, if $\delta_u > 0$,

$$\frac{1}{n^{1/2+\delta_u}} \sum_{t=1}^{n} \hat{u}_t = O_p(1).$$

Remark: Proposition 1 guarantees the use of the residuals to infer the order of summability of the error term in the model. In particular, the residuals are $S(0)$ under the null hypothesis, $\delta_u = 0$, and $S(\delta_u)$ with $\delta_u > 0$ under the alternative.

Remark: When a constant term is also estimated, that is, $y_t = \hat{m}_n + \hat{\theta}_n f(x_t) + \hat{u}_t$, where $\hat{m}_n$ is the OLS estimator of the constant term, then the OLS residuals satisfy $\sum_{t=1}^{n} \hat{u}_t = 0$, which implies that $\hat{u}_t$ cannot be used to infer δ_u. Partially demeaned residuals

$$\tilde{u}_t = \hat{u}_t - \frac{1}{t} \sum_{j=1}^{t} \hat{u}_t,$$

can be used instead in that case to infer δ_u –see BG for details on the partial demeaning procedure.

Proposition 1 motivates a way to test for strong co-summability, $H_o : \delta_u = 0$. In particular, we propose to use the order of summability estimator analyzed in BG, which is based on the convergence rate estimator by McElroy and Politis (2007). The procedure, applied to the residuals, starts from

$$S_k = \frac{1}{k^{\frac{1}{2}+\delta}} \sum_{t=1}^{k} \hat{u}_t,$$

and needs the following assumption to hold.

Assumption 2. $P(S_k = 0) = 0$ for all $k = 1, 2, 3...$

Then, following McElory and Politis (2007), the estimator of δ is based on

$$U_k = \log\left(S_k^2\right) = \log\left[\left(\frac{1}{k^{\frac{1}{2}+\delta}}\sum_{t=1}^{k}\hat{u}_t\right)^2\right],$$

which can be rewritten in regression model form

$$Y_k = \beta \log k + U_k, \quad k = 1, 2, ..., n, \tag{6}$$

where $\beta = 1 + 2\delta$ and $Y_k = \log\left[\left(\sum_{t=1}^{k}\hat{u}_t\right)^2\right]$. For expository convenience, we reproduce here the proposition in BG that derives the asymptotic distribution of $\hat{\beta}_n$, the OLS estimator of β.

Proposition 2 : *Under Assumption 2, if*

$$\frac{1}{n}\sum_{k=1}^{n} U_k \Longrightarrow D_U \quad \textit{and} \quad \frac{1}{n}\sum_{k=1}^{n}|U_k|^p = O_p(1), \tag{7}$$

for some $1 < p < \infty$ *and* D_U *a non-degenerate random variable, then*

$$\log n(\hat{\beta}_n - \beta) \Longrightarrow D_U. \tag{8}$$

Remark: To allow for a constant term in regression (6), that is,

$$Y_k = \alpha + \beta \log k + U_k,$$

where α accounts for a scaling σ in

$$U_k = \log\left(S_k^2\right) = \log\left[\left(\frac{1}{k^{\frac{1}{2}+\delta}}\frac{1}{\sigma}\sum_{t=1}^{k}\hat{u}_t\right)^2\right],$$

so that $\alpha = 2\log\sigma$, BG proposed using

$$\tilde{Y}_k = \beta \log k + \tilde{U}_k, \tag{9}$$

where $\tilde{Y}_k = (Y_k - Y_1)$ and $\tilde{U}_k = (U_k - U_1)$. In this way, inference on β is free of the nuisance parameter σ.

Remark: By Proposition 2, under the stated assumptions, the order of summability estimator $\hat{\delta}_n = \left(\hat{\beta}_n - 1\right)/2$ is consistent and when properly normalized converges to an asymptotic distribution. The approach in BG does not rely on any particular form of the data generating process; hence, some resampling technique needs to be used in order to approximate D_U and carry out inferences on the true δ. Through simulations, BG showed that the subsampling technique of Politis, Romano and Wolf (1999) worked resonably well in many situations of interest.

Next, we consider the application of this procedure to test for strong co-summability, which implies checking whether conditions (7) in Proposition 2 hold in the case of the residulas, $\hat{u}_t$. Specifically, the quantity of interest is

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}U_k &= \frac{1}{n}\sum_{k=1}^{n}\log\left[\left(\frac{1}{k^{\frac{1}{2}+\delta_u}}\sum_{t=1}^{k}\hat{u}_t\right)^2\right]\\ &= \frac{1}{n}\sum_{k=1}^{n}\log\left[\left(\frac{n^{\frac{1}{2}+\delta_u}}{k^{\frac{1}{2}+\delta_u}}\frac{1}{n^{\frac{1}{2}+\delta_u}}\sum_{t=1}^{k}\hat{u}_t\right)^2\right]\\ &= -\left(1+2\delta_u\right)\frac{1}{n}\sum_{k=1}^{n}\log\frac{k}{n}+\frac{1}{n}\sum_{k=1}^{n}\log\left[\left(\frac{1}{n^{\frac{1}{2}+\delta_u}}\sum_{t=1}^{k}\hat{u}_t\right)^2\right].\end{aligned}$$

Under the null hypothesis of strong co-summability, in which $\delta_u = 0$, for $\hat{u}_t = u_t - \left(\hat{\theta}_n - \theta_0\right) f\left(x_t\right)$,

$$\frac{1}{n}\sum_{k=1}^{n}U_k = -\frac{1}{n}\sum_{k=1}^{n}\log\frac{k}{n}+\frac{1}{n}\sum_{k=1}^{n}\log\left[\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{k}u_t - n^{\delta_f}\left(\hat{\theta}_n-\theta_0\right)\frac{1}{n^{1/2+\delta_f}}\sum_{t=1}^{k}f\left(x_t\right)\right)^2\right],$$

where, by Assumption 1 above,

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=1}^{k}\hat{u}_t &= \frac{1}{\sqrt{n}}\sum_{t=1}^{k}u_t - n^{\delta_f}\left(\hat{\theta}_n-\theta_0\right)\frac{1}{n^{1/2+\delta_f}}\sum_{t=1}^{k}f\left(x_t\right) \qquad (10)\\ &\Longrightarrow D_u\left(r\right) - D_\theta D_f\left(r\right) \equiv X\left(r\right).\end{aligned}$$

To derive the asymptotic distribution of $n^{-1}\sum_{k=1}^{n}U_k$, we will use a result by Christopeit (2009), which is, to the best of our knowledge, the most general result available in the literature that can be applied to our setup. The application of this result requires the following two assumptions to hold.

Assumption 3: $X\left(r\right)$ in (10) has continuous paths and for almost all r possesses a density ψ_r (with respect to the Lebesgue measure) such that

$$\int_0^1 \psi_r\left(x\right)dr < \infty \quad \text{for all } x.$$

Assumption 4: For each n, k, $n^{-1/2}\sum_{t=1}^{k}\hat{u}_t$ posseses a density, h_{nk}, such that

$$\sup_n \frac{1}{n}\sum_{k=1}^{n}\left\|h_{nk}\right\|_\infty < \infty.$$

Proposition 3 : *Under Assumption 1 with $\delta_u = 0$ and Assumptions 2-4,*

$$\frac{1}{n}\sum_{k=1}^{n}U_k \Longrightarrow \int_0^1 \log\left(X\left(r\right)^2\right)dr \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^{n}U_k^2 = O_p\left(1\right).$$

Remark: By Proposition 3, we can apply Proposition 2 to the least squares residuals and use $\log n(\hat{\beta}_n - 1)$ to construct a test for strong co-summability. Notice that under the null hypothesis, the relevant statistic is $\log n(\hat{\beta}_n - 1)$. Under the alternative hypothesis the appropriate quantity would be $\log n(\hat{\beta}_n - \beta_u)$ where $\beta_u = 1 + 2\delta_u > 1$. Hence, by considering $\log n(\hat{\beta}_n - 1)$ under the alternative we get a consistent test. Notice, however, that the asymptotic distribution under the null will be influenced by nuisance parameters coming from the estimation error

$$n^{\delta_f} \left(\hat{\theta}_n - \theta_0 \right) \frac{1}{n^{1/2+\delta_f}} \sum_{t=1}^{k} f(x_t) \Longrightarrow D_\theta D_f(r). \tag{11}$$

A similar feature arouse in Choi and Saikkonen (2010) when using the KPSS statistic to test for nonlinear cointegration. They proposed using a subsampling scheme to account for these nuisance parameters. In a similar fashion to Choi and Saikkonen (2010) and following BG, we propose to use the subsampling methodology of Politis, Romano and Wolf (1999) in order to carry out inferences on the order of summability of the residuals. We do not formally prove the validity of the subsampling but show through simulations that the proposal works quite well, even when compared to other tests for nonlinear cointegration existing in the literature.

In the Monte Carlo experiments, we consider the data generating processes in Table 2. Performance is measured by coverage probability of two-sided nominal 95% symmetric intervals. Size and power are measured as one minus the coverage probability that zero belongs to the corresponding subsampling confidence interval. The experiment is based on 10000 replicas and three different sample sizes, $n = \{100, 500, 1000\}$. A subsample size, $b = \sqrt{n}$, has been chosen. The results are shown in Table 3.

As it can be seen, the testing procedure is undersized –case SC, for strong co-summability– in practically all specifications being considered, except when autocorrelated errors are studied (a characteristic feature of the residual-based tests for co-integration). On the other hand, in all NSC –for no strong co-summability– cases, i.e. under the alternative, power increases as we move away from the null hypothesis (measured by the order of summability of u_t) and the sample size grows. It is worth noticing the difference between the NCS cases at the rigth and left bottom of Table 2. The DGPs on the left consider alternative hypotheses in which an additional regressor is missing in the regression function $f(x_t)$. The DGPs on the right consider cases in which the functional form has been misspecified. From results in Table 3, we see that the test has power to detect both types of misspecifications.

Table 2: DGPs: Data Generating Processes

Data Generating Processes						
*	y_t	$f(x_t)$	*		y_t	$f(x_t)$
SC	$\ln(\|x_t\|) + u_t$	$\ln(\|x_t\|)$	SC	$\rho = 0.5$	$\ln(\|x_t\|) + e_t$	$\ln(\|x_t\|)$
SC	$v_t x_t + u_t$	$v_t x_t$	SC	$\rho = 0.8$	$\ln(\|x_t\|) + e_t$	$\ln(\|x_t\|)$
SC	$x_t + v_t + u_t$	x_t	NSC	$\rho = 1$	$\ln(\|x_t\|) + e_t$	$\ln(\|x_t\|)$
SC	$x_t + 1(v_t \leq 0) x_t + u_t$	$x_t + 1(v_t \leq 0) x_t$	SC	$\rho = 0.5$	$x_t^2 + e_t$	x_t^2
SC	$x_t^2 + u_t$	x_t^2	SC	$\rho = 0.8$	$x_t^2 + e_t$	x_t^2
SC	$\sum_{j=1}^t x_j + u_t$	$\sum_{j=1}^t x_j$	NSC	$\rho = 1$	$x_t^2 + e_t$	x_t^2
*	y_t	$f(x_t)$	*		y_t	$f(x_t)$
NSC	$x_t + \ln(\|w_t\|) + u_t$	x_t	NSC		$x_t + \varepsilon_t$	$\ln(\|x_t\|)$
NSC	$x_t + v_t w_t^2 + u_t$	x_t	NSC		$x_t^2 + \varepsilon_t$	$v_t x_t$
NSC	$x_t + 1(v_t \leq 0) w_t + u_t$	x_t	NSC		$x_t + \varepsilon_t$	w_t
NSC	$x_t + w_t + u_t$	x_t	NSC		$x_t + \varepsilon_t$	x_t^2
NSC	$x_t + w_t^2 + u_t$	x_t	NSC		$x_t w_t + \varepsilon_t$	x_t^2
NSC	$x_t + \sum_{j=1}^t w_j + u_t$	x_t	NSC		$x_t + \varepsilon_t$	x_t^3

SC and *NSC* denote strong co-summability and no strong co-summability, respectively. $x_t = x_{t-1} + \varepsilon_{xt}$ with $x_0 = 0$. $w_t = w_{t-1} + \varepsilon_{wt}$ with $w_0 = 0$. $e_t = \rho e_{t-1} + z_t$. u_t, ε_t, ε_{xt}, ε_{wt}, z_t, and v_t are i.i.d. N(0,1) independent of each other.

Table 3: Testing for Strong Co-summability: Size and Power

Under Strong Co-summability											
$H_o: \delta_u = 0$			n			$H_o: \delta_u = 0$			n		
*	δ_f	δ_u	100	500	1000	*	δ_f	δ_u	100	500	1000
SC	1/2	0	0.009	0.008	0.007	SC	1/2	0	0.040	0.036	0.034
SC	1/2	0	0.011	0.013	0.011	SC	1/2	0	0.134	0.130	0.126
SC	1	0	0.008	0.010	0.010	NSC	1/2	1	0.448	0.726	0.814
SC	1	0	0.009	0.011	0.010	SC	3/2	0	0.038	0.042	0.042
SC	3/2	0	0.009	0.009	0.010	SC	3/2	0	0.146	0.153	0.155
SC	2	0	0.010	0.010	0.010	NSC	3/2	1	0.655	0.972	0.995
Under No Co-summability											
$H_o: \delta_u = 0$			n			$H_o: \delta_u = 0$			n		
*	δ_f	δ_u	100	500	1000	*	δ_y	δ_f	100	500	1000
NSC	1	1/2	0.146	0.474	0.635	NSC	1	1/2	0.166	0.490	0.682
NSC	1	1	0.213	0.569	0.742	NSC	3/2	1/2	0.946	0.999	1.000
NSC	1	1	0.207	0.612	0.792	NSC	1	1	0.510	0.919	0.979
NSC	1	1	0.511	0.920	0.978	NSC	1	3/2	0.351	0.858	0.953
NSC	1	3/2	0.701	0.981	0.998	NSC	3/2	3/2	0.706	0.994	0.999
NSC	1	2	0.769	0.982	0.997	NSC	1	2	0.448	0.894	0.961

See Table 2 for specific details about the DGPs. *SC* and *NCS* denote strong co-summability and no strong co-summability, respectively. Hence, *SC* represent size while *NCS* corresponds to power. Performance is measured from coverage probability of two-sided nominal 95% symmetric intervals.

5 Empirical Application

5.1 Asymmetric preferences of central bankers

Asymmetric preferences of central bankers have been empirically tested in the literature by using nonlinear Taylor rules. A traditional linear Taylor rule looks like

$$i_t = \theta_0 + \theta_1 \tilde{\pi}_t + \theta_3 \tilde{y}_t + u_t, \tag{12}$$

where i_t denotes nominal interest rates and $\tilde{\pi}_t$ and $\tilde{y}_t$ are deviations of inflation and output from their targets, respectively. Using equation (12), or some slightly modified version of it, several authors have tried to quantify the parameters that define the practice of monetary policy in different countries –see, for instance, Clarida, Galí and Gertler (1998, 2000).

It is somehow surprising that little attention has been paid to the fact that the variables involved in the Taylor rule are known to be highly persistent, something that should be taken into account when long time periods are analysed. There are, however, several works that address this issue, for instance, Siklos and Wohar (2005), Österholm (2005), and Christensen and Nielsen (2009). The fact that traditional Taylor rules do not appear to be congruent with the data once persistence is taken into consideration –usually through integration and co-integration theory– seems to be a common feature of these studies. This conclusion points to the possibility of an incorrect specification of the traditional Taylor rule.

On the other hand, although consistent with this conclusion, a stream of the literature has emphasised the hypothesis of asymmetric preferences of central bankers, which is often translated into non-linear Taylor rules. Next, the two cases described in Example 1 will be considered. Recall that, Clarida and Gertler (1997) consider a threshold type of Taylor rule in which the reaction of the monetary authority is different when inflation or output deviates from above, rather than from below, the target. Specifically,

$$i_t = \theta_0 + \theta_1 \tilde{\pi}_t^{(k)} 1(v_t > 0) + \theta_2 \tilde{\pi}_t^{(k)} 1(v_t \leq 0) + \theta_3 \tilde{y}_t 1(v_t > 0) + \theta_4 \tilde{y}_t 1(v_t \leq 0) + u_t, \tag{13}$$

where $\tilde{\pi}_t^{(k)}$ are deviations of the rate of inflation between periods t and $t-k$, and v_t can be either $\tilde{\pi}_t^{(k)}$ or $\tilde{y}_t$. Alternatively, Dolado, María-Dolores and Naveira (2005) derive a non-linear optimal rule when non-linearities in the Phillips curve are allowed. The main prediction of this model is that the optimal rule should contain the interaction between inflation and output gaps, that is,

$$i_t = \theta_0 + \theta_1 \tilde{\pi}_t^{(k)} + \theta_2 \tilde{y}_t + \theta_3 \tilde{\pi}_t^{(k)} \tilde{y}_t + u_t. \tag{14}$$

Note that if i_t, $\tilde{\pi}_t^{(k)}$, or $\tilde{y}_t$ are highly persistent, the non-linear nature of these two specifications invalidate the use of standard co-integration theory to analyse the relevance of these models. Nevertheless, co-summability can be used given its generality when allowing for persistence and non-linearities at the same time. Moreover, the linearity in parameters of both equations makes it suitable the application of the tools to test for co-summability proposed above.

To this end, we use US monthly time series covering the period 1954:07-2013:03, which are obtained from the Federal Reserve Bank of St. Louis. Specifically, we use (i) federal funds rate as interest rates, (ii) annual ($t/t-12$ basis; $k = 12$) percentage rate in the CPI for inflation, (iii) (logged) industrial production index for output. Following the usual practice in the literature, to measure the output gap, we detrend (logged) industrial production using the HP filter with a coefficient of 14.800. For the inflation target, we use a fixed 2% level. Figure 1 shows the temporal evolution of these three measures $-i_t$, $\tilde{\pi}_t^{(k)}$, and $\tilde{y}_t$.

Figure 1: Optimal Rules of Monetary Policy

Variables: Taylor Rule

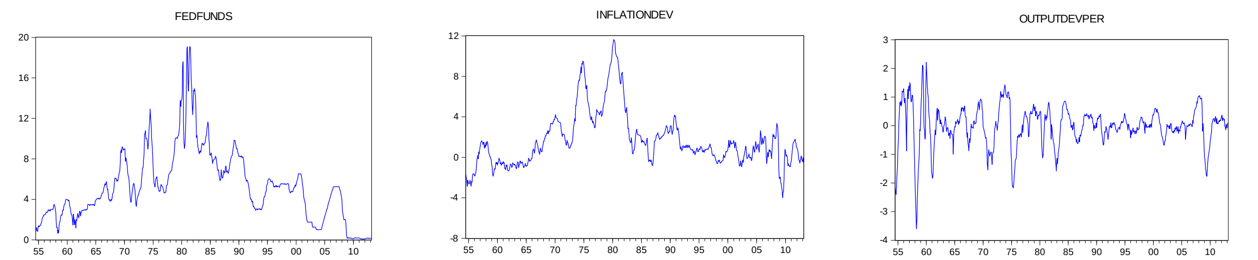


Table 4 reports the estimated orders of summability of all the variables contained in equations (13) and (14) as well as their corresponding subsampling confidence interval. All the variables have been partially demeaned to compute their orders of summability. Moreover, to control for a possible constant term in regression model (6) the first observation is subtracted –that is, equation (9) has been used.

Table 4: Order of Summability: Estimation and Inference

Variables	$\hat{\delta}$	I_L	I_U
i_t	0.813	0.419	1.207
$\tilde{\pi}_t^{(k)}$	0.862	0.404	1.321
$\tilde{y}_t$	0.490	0.055	0.925
$\tilde{\pi}_t^{(k)}\tilde{y}_t$	0.198	-0.381	0.778
$\tilde{\pi}_t^{(k)}1\left(\tilde{\pi}_t^{(k)} > 0\right)$	0.814	0.459	1.169
$\tilde{\pi}_t^{(k)}1\left(\tilde{\pi}_t^{(k)} \leq 0\right)$	0.697	0.271	1.122
$\tilde{y}_t1\left(\tilde{\pi}_t^{(k)} > 0\right)$	0.155	-0.502	0.813
$\tilde{y}_t1\left(\tilde{\pi}_t^{(k)} \leq 0\right)$	0.398	-0.232	1.029
$\tilde{\pi}_t^{(k)}1\left(\tilde{y}_t > 0\right)$	0.805	0.301	1.309
$\tilde{\pi}_t^{(k)}1\left(\tilde{y}_t \leq 0\right)$	0.725	0.240	1.210
$\tilde{y}_t1\left(\tilde{y}_t > 0\right)$	0.496	0.129	0.862
$\tilde{y}_t1\left(\tilde{y}_t \leq 0\right)$	0.626	0.186	1.065

$\hat{\delta}$ denotes the estimated order of summability calculated from regression (10). I_L and I_U denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals. All the variables have been partially demeaned.

Results in Table 4 indicate that interest rates, i_t, and inflation gap, $\tilde{\pi}_t^{(k)}$, have a similar order of summability of approximately 0.8, while the estimated order of summability for the output gap, $\tilde{y}_t$, is approximately 0.5. It is worth emphasising that zero does not belong to any of the subsampling confidence intervals of these three time series. This confirms that persistence has to be properly addressed when using this dataset. With respect to the non-linear variables, different results are found. While the subsampling confidence intervals for $\tilde{\pi}_t^{(k)}\tilde{y}_t$, $\tilde{y}_t1\left(\tilde{\pi}_t^{(k)} > 0\right)$, and $\tilde{y}_t1\left(\tilde{\pi}_t^{(k)} \leq 0\right)$ do contain zero, all the others do not.

Table 5 collects the parameter estimates of equations (13) and (14) jointly with the results of testing for strong co-summability associated with each regression. Some aspects are worth emphasising. First, the traditional Taylor rule does not specify a strong co-summable relationship –zero does not belong to the corresponding subsampling confidence interval. Second, focusing on the non-linear specifications, it can be seen that only a threshold type of Taylor rule in which the Federal Reserve reacts asymmetrically to output deviations is not rejected –zero belongs to the interval in this case. Finally, the difference between the parameters associated to $\tilde{y}_t1\left(\tilde{y}_t > 0\right)$ and $\tilde{y}_t1\left(\tilde{y}_t \leq 0\right)$ is remarkable. This fact clearly reflects a greater aversion to recessions than to expansions of the monetary authorities in the US.

Table 5: Testing for Co-summability

Taylor Rules	i_t	i_t	i_t	i_t
1	3.641	3.606	3.603	3.781
$\tilde{\pi}_t^{(k)}$	0.957	0.959		
$\tilde{y}_t$	0.744	0.445		
$\tilde{\pi}_t^{(k)}\tilde{y}_t$		0.173		
$\tilde{\pi}_t^{(k)}1\left(\tilde{\pi}_t^{(k)} > 0\right)$			0.965	
$\tilde{\pi}_t^{(k)}1\left(\tilde{\pi}_t^{(k)} \leq 0\right)$			0.916	
$\tilde{y}_t1\left(\tilde{\pi}_t^{(k)} > 0\right)$			0.974	
$\tilde{y}_t1\left(\tilde{\pi}_t^{(k)} \leq 0\right)$			0.290	
$\tilde{\pi}_t^{(k)}1\left(\tilde{y}_t > 0\right)$				1.052
$\tilde{\pi}_t^{(k)}1\left(\tilde{y}_t \leq 0\right)$				0.820
$\tilde{y}_t1\left(\tilde{y}_t > 0\right)$				0.119
$\tilde{y}_t1\left(\tilde{y}_t \leq 0\right)$				0.807
$\hat{\delta}_{\hat{u}}$	0.428	0.471	0.437	0.403
I_L	0.036	0.087	0.011	-0.005
I_U	0.819	0.854	0.863	0.811

$\hat{\delta}_{\hat{u}}$ denotes the estimated order of summability of the residuals calculated from regression (10). Residuals have been partially demeaned. I_L and I_U denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals.

5.2 Environmental Kuznets Curve

The environmental Kuznets curve –EKC– suggests an inverted U-shaped relationship between pollution and economic development –see Dasgupta et al. (2001), Grossmann and Krueger (1995) or Brock and Taylor (2005). The empirical literature on the EKC has mainly used a reduced form approach. Typically, polynomial relationships between pollution and income have been considered, that is,

$$p_t = \theta_0 + \theta_1 y_t + \theta_2 y_t^2 + ... + \theta_k y_t^k + u_t, \tag{15}$$

where p_t is a measure of pollution and y_t is a measure of income. Several empirical issues arise in this setup. A first issue is concerned with the measures chosen for p_t and y_t. While GDP has been used as a measure of income many measures of pollutants have been used. Commonly used measures for p_t are CO_2, NO_x, and SO_2. Empirical evidence is mixed for different pollutants. A second issue relates to the curvature of the EKC. There seems not to be a clear agreement about the order of the polynomial to be used. Grossman and Krueger (1995) used a cubic specification, while Holtz-Eakin and Selden (1995) considered the quadratic one. Other authors tend to compare both specifications in practice. A third empirical ambiguity arises as p_t and y_t are sometimes treated in levels (Grossman and Krueger, 1995), other times in logarithms (Hong and Wagner, 2008) or both cases are compared (Holtz-Eakin and Selden, 1995). Finally, it is surprising that only a few authors have taken into consideration persistence of the variables involved in the EKC. Some exceptions include Perman and Stern (2003), Hong and Wagner (2008) and Jalil and Mahmud (2009). When persistence is taken into consideration, the empirical evidence on the EKC is mixed.

As an illustration, we apply co-summability theory to disentangle some of the empirical features on the EKC. We use annual GDP and CO_2 emissions per capita in the US during the period 1870-2007. GDP and population are taken from Angus Maddison and CO_2 emissions from the Carbon Dioxide Information Analysis Centre. Figures 2-3 show the evolution of GDP and CO_2 emissions per capita, in levels –$co2pcus$, $gdppcus$– and logarithms –$lco2pcus$, $lgdppcus$.

Figure 2: Environmental Kuznets Curve Hypothesis

Variables in levels: Environmental Kuznets Curve

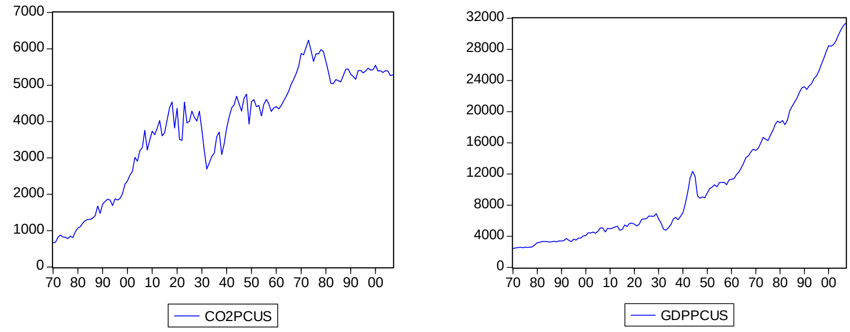


Figure 3: Environmental Kuznets Curve Hypothesis

Variables in Logs: Environmental Kuznets Curve

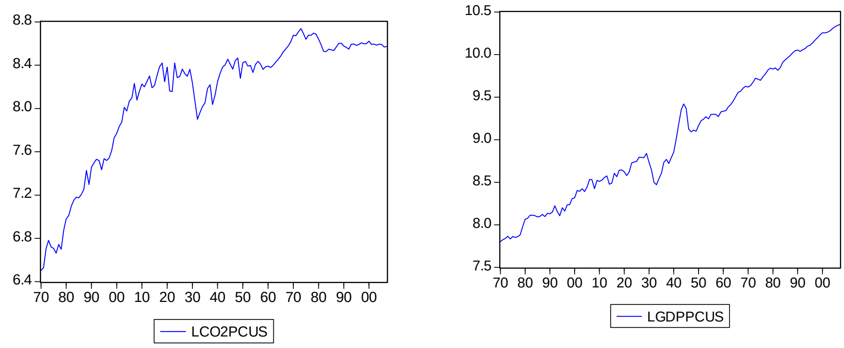


Table 6 reports the estimated orders of summability of all the variables contained in (15) for $k = 4$. The corresponding subsampling confidence intervals are provided as well. As expected, the order of summability of GDP per capita increases as successive powers are taken. In general, these results show that persistence must be taken into account –zero does not belong to any of the confidence intervals.

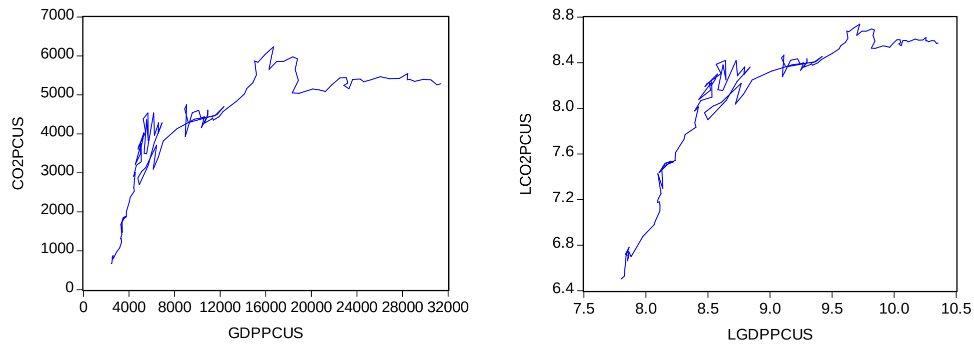
Table 6: Order of Summability: Estimation and Inference

Variables	$\hat{\delta}$	I_L	I_U
$co2pc$	0.893	0.286	1.500
$gdppc$	1.424	0.599	2.249
$gdppc^2$	1.779	0.795	2.763
$gdppc^3$	2.090	0.952	3.229
$gdppc^4$	2.391	1.082	3.699
$lco2pc$	0.705	0.160	1.250
$lgdppc$	0.876	0.195	1.557
$lgdppc^2$	0.950	0.255	1.645
$lgdppc^3$	1.017	0.270	1.764
$lgdppc^4$	1.112	0.260	1.963

$\hat{\delta}$ denotes the estimated order of summability calculated from regression (10). I_L and I_U denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals. All the variables have been partially detrended.

Figure 4 plots the relationship between GDP and CO_2 emissions per capita in levels and logarithms. Although it seems there is a diminishing marginal propensity to pollute, the postulated inverted U-shape should be more carefully and formally tested.

Figure 4: Environmental Kuznets Curve Hypothesis



Results to test for strong co-summability are collected in Tables 7 and 8 for levels and logarithms, respectively. From Table 7, it is clear that co-summability does not hold in any of the specifications considered. Nevertheless, in Table 8, results are somehow more optimistic. Co-summability is not rejected for the cubic specification, which is compatible with the shape observed in Figure 4. These results are invariant to the inclusion of a deterministic trend, a usual practice in the literature. Summarising, from the co-summability results, we recommend to use the logarithmic transformation and polynomials of third degree when empirically studying parametric reduced forms of the EKC in the US from this database.

Table 7: Testing for Co-summability

EKC	$co2$	$co2$	$co2$	$co2$	$co2$	$co2$
1	2190.538	1090.780	470.265	749.636	-1040.9460	-391.214
t		54.820		39.204		20.297
gdp	0.149	-0.098	0.520	0.112	1.0190264	0.641
gdp^2			-1.249e-005	-4.752e-006	-5.143e-005	-3.441e-005
gdp^3					8.248e-010	5.491e-010
$\hat{\delta}_{\hat{u}}$	1.750	0.883	1.447	0.989	1.028	0.923
I_L	0.945	0.143	0.483	0.322	0.272	0.229
I_U	2.556	1.623	2.411	1.656	1.785	1.617

$\hat{\delta}_{\hat{u}}$ denotes the estimated order of summability of the residuals calculated from regression (10). Residuals have been partially demeaned. I_L and I_U denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals.

Table 8: Testing for Co-summability

EKC	$lco2$	$lco2$	$lco2$	$lco2$	$lco2$	$lco2$
1	2.290	10.023	-42.883	-41.941	-280.718	-290.421
t		0.019		0.001		-0.003
$lgdp$	0.646	-0.359	10.665	10.501	90.013	92.866
$lgdp^2$			-0.551	-0.546	-9.340	-9.623
$lgdp^3$					0.323	0.333
$\hat{\delta}_{\hat{u}}$	1.503	1.351	0.792	0.796	0.240	0.247
I_L	0.724	0.529	0.189	0.172	-0.342	-0.305
I_U	2.281	2.172	1.395	1.419	0.823	0.801

$\hat{\delta}_{\hat{u}}$ denotes the estimated order of summability of the residuals calculated from regression (10). Residuals have been partially demeaned. I_L and I_U denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals.

6 Concluding Remarks

Co-integration theory is not designed to deal with situations in which non-linearities and persistence occur simultaneously. Accordingly, there is a clear need for theoretically valid and empirically useful concepts that generalise those of integration and co-integration to non-linear environments. To that end, this paper has made use of the order of summability of stochastic processes to formalize the idea of co-summability, a generalization of co-integration for nonlinear relationships, and has developed a simple testing procedure to analyze it in practice.

7 Appendix

Proof of Proposition 1:

(a) Under Assumptions 1, if $\delta_u = 0$,

$$\begin{aligned}\frac{1}{n^{1/2}}\sum_{t=1}^{n}\hat{u}_t &= \frac{1}{n^{1/2}}\sum_{t=1}^{n}\left(y_t - \hat{\theta}_n f(x_t)\right)\\ &= \frac{1}{n^{1/2}}\sum_{t=1}^{n}\left(\theta_0 f(x_t) + u_t - \hat{\theta}_n f(x_t)\right)\\ &= \frac{1}{n^{1/2}}\sum_{t=1}^{n} u_t - n^{\delta_f}\left(\hat{\theta}_n - \theta_0\right)\frac{1}{n^{1/2+\delta_f}}\sum_{t=1}^{n} f(x_t)\\ &= O_p(1).\end{aligned}$$

(b) Under Assumption 1, if $\delta_u > 0$,

$$\begin{aligned}\frac{1}{n^{1/2+\delta_u}}\sum_{t=1}^{n}\hat{u}_t &= \frac{1}{n^{1/2+\delta_u}}\sum_{t=1}^{n}\left(\theta_0 f(x_t) + u_t - \hat{\theta}_n f(x_t)\right)\\ &= \frac{1}{n^{1/2+\delta_u}}\sum_{t=1}^{n} u_t + n^{\delta_f-\delta_u}\left(\hat{\theta}_n - \theta_0\right)\frac{1}{n^{1/2+\delta_f}}\sum_{t=1}^{n} f(x_t)\\ &= O_p(1).\end{aligned}$$

Q.E.D.

Proof of Proposition 3: For $\hat{u}_t = u_t - \left(\hat{\theta}_n - \theta_0\right) f(x_t)$ with $\delta_u = 0$, we have

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} U_k &= \frac{1}{n}\sum_{k=1}^{n}\log\left[\left(\frac{1}{k^{\frac{1}{2}}}\sum_{t=1}^{k}\hat{u}_t\right)^2\right]\\ &= -\frac{1}{n}\sum_{k=1}^{n}\log\frac{k}{n} + \frac{1}{n}\sum_{k=1}^{n}\log\left[\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{k} u_t - n^{\delta_f}\left(\hat{\theta}_n - \theta_0\right)\frac{1}{n^{1/2+\delta_f}}\sum_{t=1}^{k} f(x_t)\right)^2\right].\end{aligned}$$

By Assumption 1,

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=1}^{k}\hat{u}_t &= \frac{1}{\sqrt{n}}\sum_{t=1}^{k} u_t - n^{\delta_f}\left(\hat{\theta}_n - \theta_0\right)\frac{1}{n^{1/2+\delta_f}}\sum_{t=1}^{k} f(x_t)\\ &\Longrightarrow D_u(r) - D_\theta D_f(r) \equiv X(r).\end{aligned}$$

Now, under Assumptions 2-4, Theorem 2.1. in Christopeit (2009) jointly with an extended real functions argument that follows Remark 2.1 in Pötscher (2004),

$$\frac{1}{n}\sum_{k=1}^{n} U_k \Longrightarrow \int_0^1 \log\left(X(r)^2\right) dr.$$

By similar arguments,

$$\frac{1}{n}\sum_{k=1}^{n}U_k^2 \Longrightarrow \int_0^1 \log^2\left(X\left(r\right)^2\right)dr,$$

hence,

$$\frac{1}{n}\sum_{k=1}^{n}U_k^2 = O_p\left(1\right).$$

Q.E.D.

8 References

Berenguer-Rico, V., and J. Gonzalo (2014): "Summability of stochastic processes: A generalization of integration for non-linear processes," *Journal of Econometrics*, 178, 331-341.

Brock, W. A., and M. S. Taylor (2005): "Economic growth and the environment: A review of theory and empirics," *Handbook of Economic Growth*, Volume 1, Part 2, 1749-1821.

Choi, I., and P. Saikkonen (2010): "Tests for non-linear cointegration," *Econometric Theory*, 26, 682-709.

Christensen, A. M., and H. B. Nielsen (2009): "Monetary Policy in the Greenspan Era: A Time Series Analysis of Rules vs. Discretion," *Oxford Bulletin of Economics and Statistics*, 71, 69-89.

Christopeit, N. (2009): "Weak Convergence of Nonlinear Transformations of Integrated Processes: The Multivariate Case," *Econometric Theory*, 25, 1180-1207.

Clarida, R., and M. Gertler (1997): "How the Bundesbank conducts monetary policy." In: Romer, D. (Ed.), *Reducing Inflation*. Chicago: University of Chicago Press.

Clarida, R., J. Galí, and M. Gertler (1998): "Monetary policy rules in practice. Some international evidence," *European Economic Review*, 42, 1033-1067.

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

Dasgupta, S., B. Laplante, H. Wang, and D. Wheeler (2002): "Confronting the Environmental Kuznets Curve," *The Journal of Economic Perspectives*, 16, 147-168.

Dolado, J. J., R. María-Dolores, and M. Naveira (2005): "Are monetary-policy reaction functions asymmetric?: The role of non-linearity in the Phillips curve," *European Economic Review*, 49, 485-503.

Embrechts, P., C. Klüppelberg, and T. Mikosh (1999): *Modelling extremal events for insurance and finance*. Berlin: Springer-Verlag.

Gonzalo, J., and J. Y. Pitarakis (2006): "Threshold effects in co-integrating regressions," *Oxford Bulletin of Economics & Statistics*, 68, 813-833.

Granger, C. W. J. (1995): "Modelling non-linear relationships between extended-memory variables," *Econometrica*, 63, 265-279.

Grossman, G. M., and A. B. Krueger (1995): “Economic growth and environment,” *Quarterly Journal of Economics*, 110, 353-377.

Holtz-Eakin, D., and T. M. Selden (1995): “Stoking the fires? CO_2 emissions and economic growth,” *Journal of Public Economics*, 57, 85-101.

Hong, S. H., and M. Wagner (2008): “Non-linear cointegration analysis and the environmental Kuznets curve,” Mimeo.

Jalil, A., and S. F. Mahmud (2009): “Environment Kuznets curve for CO_2 emissions: A cointegration analysis for China,” *Energy Policy*, 37, 5167-5172.

Karlsen, H. A., T. Myklebust, and D. Tjϕstheim (2007): “Nonparametric estimation in a nonlinear cointegration type model,” *The Annals of Statistics*, 35, 252-299.

Kasparis, I. (2008): “Detection of functional form misspecification in cointegrating relations,” *Econometric Theory*, 24, 1373-1403.

McElroy, T., and D. N. Politis (2007): “Computer-intensive rate estimation, diverging statistics, and scanning,” *The Annals of Statistics*, 35, 1827-1848.

Österholm, P. (2005): “The Taylor Rule: A Spurious Regression?,” *Bulletin of Economic Research*, 57, 217-247.

Park, J. Y., and P. C. B. Phillips (1999): “Asymptotics for non-linear transformations of integrated time series,” *Econometric Theory*, 15, 269-298.

Park, J. Y., and P. C. B. Phillips (2001): “Non-linear regressions with integrated time series,” *Econometrica*, 69, 117-161.

Perman, R. J., and D. Stern (2003): “Evidence from panel unit root and cointegration tests that the Environmental Kuznets Curve does not exist,” *Australian Journal of Agricultural and Resource Economics*, 47, 325-347.

Phillips, P. C. B. (1986): “Understanding Spurious Regressions in Econometrics,” *Journal of Econometrics*, 33, 311-340.

Politis, D.N., J.P. Romano, and M. Wolf (1999): “Subsampling.” *Springer*, New York.

Pötscher, B. M. (2004): “Non-linear functions and convergence to Brownian motion: beyond the continuous mapping theorem,” *Econometric Theory,* 20, 1-22.

Schienle, M. (2011): “Nonparametric nonstationary regression with many covariates,” Mimeo.

Siklos, P. L., and M. E. Wohar (2005): “Estimating Taylor-type rules: An unbalanced regression?,” *Econometric Analysis of Financial and Economic Time Series*, 20, 239-276.

Wang, Q., and P. C. B. Phillips (2009): “Structural nonparametric cointegrating regressions,” *Econometrica*, 77, 1901-1948.