Documentos de Trabajo

Lan Gaiak

ESTIMATION OF THE COINTEGRATING RANK IN FRACTIONAL COINTEGRATION

Javier Hualde

D.T. 1205

Departamento de Economía

Ekonomia Saila

Universidad Pública de Navarra

Nafarroako Unibertsitate Publikoa

Estimation of the cointegrating rank in fractional cointegration*

Javier Hualde
Universidad Pública de Navarra

August 9, 2012

Abstract

This paper proposes an estimator of the cointegrating rank of a potentially cointegrated multivariate fractional process. Our setting is very flexible, allowing the individual observable processes to have different integration orders. The proposed method is automatic and can be also employed to infer the dimensions of possible cointegrating subspaces, which are characterized by special directions in the cointegrating space which generate cointegrating errors with smaller integration orders, increasing the "achievement" of the cointegration analysis. A Monte Carlo experiment of finite sample performance and an empirical analysis are included.

JEL Classification: C32.

Keywords. Fractional integration; cointegrating rank; cointegrating space and subspaces.

*Thanks to Søren Johansen for useful comments. This research is supported by the Spanish Ministerio de Ciencia e Innovación (ref. ECO2008-02641).

1. INTRODUCTION

Since the seminal paper of Engle and Granger (1987), the concept of cointegration has been generalized in several directions. One of the most recent developments is that of fractional cointegration, where, unlike in the standard setting of unit root observables with weak dependent cointegrating errors, series are allowed to have fractional integration orders. This generalization captures interesting possibilities, like that of nonstationary but mean reverting cointegration errors (with applications in macroeconomics), or stationary cointegration, where the observables are stationary (but long memory) with "less-memoried" (even short memory) cointegrating errors, with applications in finance (see Gil-Alana and Hualde, 2009, for a review).

However, even if theoretical developments have been numerous, it is not clear the extent to which they are influencing the empirical literature, possibly due to the absence of a feasible and general methodology. In particular, even if most studies impose that all observables share the same integration order (Kim and Phillips, 2000, Hurvich and Chen, 2006), it seems natural to think that this restriction does not apply in practice. Then, a complicated cointegrating structure might occur, and issues like determining the cointegrating rank r (the dimension of the cointegrating space), which is necessary in order to make inference on cointegrating vectors, are far from being trivial. In fact, methodologies for estimating the rank, as the ones proposed by Robinson and Yajima (2002) or Robinson (2008), are not designed to cover all cases where observables have distinct integration orders. Thus, our purpose in the present paper will be to propose an automatic method to infer r which does not require any prior information about memory or cointegrating characteristics of the observables. Incidentally, our method offers two additional advantages. First, it leads to straightforward estimation of the cointegrating space. Second, our methodology could be also used to uncover the dimension of possible cointegrating subspaces (special directions in the cointegrating space which generate cointegrating errors with smaller integration orders), and estimate them by simple methods. In this sense, our procedure could be viewed as an alternative to Chen and Hurvich (2006), although they assume that all observables share the same memory parameter and, in addition, our approach is much closer to the simultaneous equations model methodology, with very long tradition in econometrics.

The plan of the paper is as follows. In Section 2 we propose an estimator of the dimension of the cointegrating space and give guidance on how to estimate it. Next, in Section 3, we present a Monte Carlo experiment of finite sample performance. In Section 4 we apply our method to a trivariate series of oil prices, and illustrate the issue of estimating a cointegrating subspace. Finally, in Section 5, we conclude. All proofs are relegated to the Appendix.

2. ESTIMATION OF THE COINTEGRATING RANK

First, we model a $p \times 1$ vector of observables z_t. Let u_t be a p-dimensional covariance stationary process with spectral density positive definite and bounded at all frequencies; for real numbers θ_i, $i = 1, ..., p$, such that $\theta_i < 1/2$, define $s_t = diag\left(\Delta^{-\theta_1}, ..., \Delta^{-\theta_p}\right) u_t$, where $\Delta = 1 - L$, L being the lag operator. Denote by a_{it} the ith component of an arbitrary vector a_t. Then $s_{it} \sim I(\theta_i)$, where $I(d)$ stands for Type I fractionally integrated process of order d (see, e.g., Gil-Alana and Hualde, 2009). Let $q_i \in \{0, 1, 2, ...\}$, $i = 1, ..., p$. Define $v_t = diag\left(\Delta^{-q_1}, ..., \Delta^{-q_p}\right)\{s_t 1(t > 0)\}$, where $1(\cdot)$ is the indicator function (this implies fixing initial conditions to zero, but our results are identical for initial conditions fixed in alternative ways). Let $\delta_i = \theta_i + q_i$. Then $v_{it} \sim I(\delta_i)$, which is covariance stationary if $q_i = 0$, or built on partial sums of a covariance stationary process if $q_i > 0$. Without loss of generality (apart from not allowing negative memories), set $0 \leq \delta_1 \leq \delta_2 \leq ... \leq \delta_p$, with also $\delta_p > 0$. Let the $p \times 1$ vector of observables z_t be generated by

$$\Upsilon z_t = v_t, \; t = 1, 2, ..., \tag{1}$$

where Υ is nonsingular. Then, $z_t = \Upsilon^{-1} v_t$, so the variables in z_t are modelled as linear combinations of fractional processes (hence, inheriting, in general, the maximum of their integration orders). Model (1) appears to be a natural setting to discuss the possibility of cointegration, which throughout will refer to the situation where a linear combination of fractional processes is integrated of a strictly smaller order than the maximum order of the elements of the linear combination. This definition covers many situations. For example, if one of the variables has an integration order strictly larger than the rest of the variables, any linear combination which puts zero weight on this particular variable is considered to be a (trivial) cointegrating relation. The definition is similar to that of Johansen (1995), and more general than those of Flores and Szafarz (1996), Marinucci and

Robinson (2001) and Robinson and Yajima (2002). When all variables enjoy the same integration order, all different definitions coincide, being also identical to the original one given in Engle and Granger (1987).

Related to (1), the nonsingularity of Υ and the assumptions on u_t (so nontrivial cointegration cannot occur among the elements of v_t), imply that at least one of the elements of z_t is $I(\delta_p)$, although the possibility that all variables in z_t share the same integration order is also captured if the last column of Υ^{-1} contains no zeroes. In general $z_{it} \sim I(d_i)$, $i = 1, ..., p$, where the linkage between the d_i's and the δ_i's depends on Υ. Note that by (1), z_t might be subject to a very complicated cointegrating structure (depending both on restrictions in Υ, which might eliminate trends in certain linear combinations, and strict inequalities among the δ_i's)

The main aim of the paper is to propose an estimator of the cointegrating rank of z_t. Our methodology will be based on applying sequentially the following theorem. We do not give its proof, as it is just Theorem 2 of Gomez-Biscarri and Hualde (2011) applied to our fractional setting. Denote by "common trends" $I(\delta_p)$ and noncointegrated variables.

Theorem 1. z_t in (1) has cointegrating rank $r \in \{1, ..., p-1\}$ if and only if a. and b. hold, where: a. There exists a $(p-r)$-dimensional subvector of z_t (say $z_{(b)t}$), whose individual components are common trends; b. All subvectors of z_t of dimension larger than $p - r$ containing $z_{(b)t}$ cointegrate.

The implementation of our procedure is based on estimators of the individual integration orders ($\widehat{d}_i$, $i = 1, ..., p$) and test statistics ($\widehat{\tau}_{j_1,...,j_k}$) for

$$H_{j_1,...,j_k} : \{z_{j_1t}, z_{j_2t}, ..., z_{j_kt} \text{ are not cointegrated}\}\,; \overline{H}_{j_1,...,j_k} : H_{j_1,...,j_k} \text{ is not true},$$

where $j_1, ..., j_k \in \{1, ..., p\}$, $k \leq p$. Let g_{1n} be the rate of convergence of the $\widehat{d}_i$'s, and g_{2n} be the rate of divergence under the alternative of the $\widehat{\tau}_{j_1,...,j_k}$'s (if it is different for the various test statistics, take g_{2n} as the minimum of the rates). Also, assume that $\widehat{\tau}_{j_1,...,j_k}$ has asymptotic size α. Possible choices for $\widehat{d}_i$ are the log-periodogram (Robinson, 1995a) or the local Whittle (Robinson, 1995b) estimators, for which $g_{1n} = m^{1/2}$, where m denotes bandwidth. Regarding $\widehat{\tau}_{j_1,...,j_k}$, Robinson (2008) (where $g_{2n} = m$) or Hualde and Velasco (2008) (where g_{2n} depends on a complicated manner on the integration orders involved, m and n) are good alternatives, because they do not impose equality of integration orders of the observables. Let $h_n > 0$ be a sequence (whose role will be clarified in Theorem 2 and Remark 1 below) such that

$$h_n + (g_{1n} + g_{2n})\, h_n^{-1} \to \infty \text{ as } n \to \infty. \tag{2}$$

The steps characterizing our procedure are as follows:

Step 1. Estimate d_i by $\widehat{d}_i$, $i = 1, ..., p$. Then choose a possible common trend variable $z_{c_1 t}$, $c_1 \in \{1, ..., p\}$. In particular, choose $c_1 = p$ if $g_{1n}(\widehat{d}_p - \widehat{d}_i) > h_n$, for all $i = 1, ..., p-1$; otherwise, choose $c_1 = p - j$ if $g_{1n}(\widehat{d}_{p-j} - \widehat{d}_i) > h_n$, for all $i = 1, ..., p-j-1$, checking these conditions sequentially for $j = 1, ..., p-2$; if none of these conditions is fulfilled, choose $c_1 = 1$. Note that there is a particular ordering in our choice, but the limiting properties of our proposed estimator of r are invariant to it. Next, reorder the variables in z_t so that $z_{pt} = z_{c_1 t}$ in the new ordering (this reordering is irrelevant for the results, but simplifies subsequent notation substantially). Then, given the possible common trend z_{pt}, we test for $H\,(1) : \cup_{i=1}^{p-1} H_{p,i}$, $\overline{H}\,(1) : \cap_{i=1}^{p-1} \overline{H}_{p,i}$, and set $\{\widehat{r} = p-1\} = \{H\,(1) \text{ is rejected}\}$. Note that if $z_{pt} \sim I\,(\delta_p)$, then by Theorem 1, $H\,(1)$, $\overline{H}\,(1)$ are equivalent to $r < p-1$, $r = p-1$, respectively, which justifies $\widehat{r}$.

Step 2. If $H\,(1)$ is not rejected, choose $z_{c_2 t}$, $c_2 \in \{1, ..., p-1\}$ so the possible common trends are z_{pt}, $z_{c_2 t}$. In particular, choose $c_2 = p-1$ if $\widehat{\tau}_{p,i} - \widehat{\tau}_{p,p-1} > h_n$ for all $i = 1, ..., p-2$; otherwise, choose $c_2 = p-j$ if $\widehat{\tau}_{p,i} - \widehat{\tau}_{p,p-j} > h_n$ for all $i = 1, ..., p-j-1$, checking these conditions sequentially for $j = 2, ..., p-2$; if none of these conditions is fulfilled, choose $c_2 = 1$. Reorder again the variables so that $z_{pt} = z_{c_1 t}$, $z_{p-1,t} = z_{c_2 t}$ in the new ordering. Then we test for $H\,(2) : \cup_{i=1}^{p-2} H_{p,p-1,i}$, $\overline{H}\,(2) : \cap_{i=1}^{p-2} \overline{H}_{p,p-1,i}$, and estimate the rank by

$$\{\widehat{r} = p-2\} = \{H\,(1) \text{ is not rejected and } H\,(2) \text{ is rejected}\}.$$

Note that if z_{pt}, $z_{p-1,t}$ are valid common trends, then $H\,(1) \cap H\,(2)$, $H\,(1) \cap \overline{H}\,(2)$ are equivalent to $r < p-2$, $r = p-2$, respectively, which justifies $\widehat{r}$.

In general, for $k = 2, ..., p-1$, we have

Step k. If $H\,(k-1)$ is not rejected, choose c_k. Note that in previous steps the variables have been reordered so that $z_{pt} = z_{c_1 t}, ..., z_{p-k+2,t} = z_{c_{k-1},t}$. Then choose $c_k = p-k+1$ if $\widehat{\tau}_{p,...,p-k+2,i} - \widehat{\tau}_{p,...,p-k+2,p-k+1} > h_n$ for all $i = 1, ..., p-k$; otherwise, choose $c_k = p-j$ if $\widehat{\tau}_{p,...,p-k+2,i} - \widehat{\tau}_{p,...,p-k+2,p-j} > h_n$ for all

$i = 1, ..., p - j - 1$, checking these conditions sequentially for $j = k, ..., p - 2$; if none of these conditions is fulfilled, choose $c_k = 1$. Reorder the variables so $z_{pt} = z_{c_1 t}, ..., z_{p-k+2,t} = z_{c_{k-1},t}, z_{p-k+1,t} = z_{c_k t}$. Then test for $H(k) : \cup_{i=1}^{p-k} H_{p,p-1,...,p-k+1,i}$, $\overline{H}(k) : \cap_{i=1}^{p-k} \overline{H}_{p,p-1,...,p-k+1,i}$, and set

$$\{\widehat{r} = p - k\} = \{H(i), \, i = 1, ..., k-1, \text{ are not rejected and } H(k) \text{ is rejected}\},$$

and for the step $k = p-1$, also $\{\widehat{r} = 0\} = \{H(i), \, i = 1, ..., p-1, \text{ are not rejected}\}$.

Before analyzing the properties of $\widehat{r}$, we derive a result concerning our choice of common trends. We introduce first some additional notation. Given the initial arbitrary ordering of the variables, let $i_1 \in \{1, ..., p\}$ be such that: $z_{i_1 t} \sim I(\delta_p)$; $i_1 \leq l$ for any l such that $z_{lt} \sim I(\delta_p)$. Similarly, given the ordering of z_t after c_1 has been chosen (so $z_{pt} = z_{c_1 t}$), let $i_2 \in \{1, ..., p-1\}$ be such that: z_{pt} and $z_{i_2 t}$ are not cointegrated; $i_2 \leq l$ for any l such that z_{pt} and z_{lt} are not cointegrated. In general, for $j = 2, ..., p-1$, let $i_j \in \{1, ..., p-j+1\}$ be such that: $z_{pt}, ..., z_{p-j+2,t}$ and $z_{i_j t}$ are not cointegrated; $i_j \leq l$ for any l such that $z_{pt}, ..., z_{p-j+2,t}$and z_{lt} are not cointegrated. Note that the existence of the i_j's depends on r. In particular, if $r = p-1$, just i_1 exists; if $r = p-2$, just i_1 and i_2 exist; in general, for $r \in \{1, ..., p-1\}$, just $i_1, i_2, ..., i_{p-r}$ exist.

Theorem 2. Let $r \in \{1, ..., p-1\}$ be the cointegrating rank and (2) hold. Then, for any $1 \leq k \leq p - r$, $\Pr(c_1 = i_1, c_2 = i_2, ..., c_k = i_k) \to 1$ as $n \to \infty$.

Remark 1. Theorem 2 implies that, by using h_n such that (2) holds, we choose in every step a single set of valid common trends with probability approaching one. Alternatively, it would be more natural to set $h_n = 0$, because this implies that $z_{c_1 t}$ is the variable with highest estimated order, $z_{c_2 t}$ would be the variable which shows less evidence of being cointegrated with $z_{c_1 t}$, and so on. However, in this case, alternative sets of valid common trends could be chosen with nonnegligible (as $n \to \infty$) probabilities (unlike in our setting, where just a particular set of valid common trends has a nonnegligible probability of being chosen), and this leads to a size control problem. Using h_n satisfying (2) implies a unique choice of valid common trends (asymptotically), and this allows us to control the size of our sequential method and derive the neat results (3), (4), (5), (6) below. In practice, however, there is always a choice for h_n as close to zero as desired, while satisfying (2) (as the one we employ in the Monte Carlo experiment).

The properties of $\widehat{r}$ are given in the next theorem.

Theorem 3. Let r be the cointegrating rank and (2) hold. Then, as $n \to \infty$,

$$\Pr(\widehat{r} = j) \to 0, \quad j = 0, ..., r-1, \tag{3}$$

$$\Pr(\widehat{r} = j) \leq \phi_n \to \alpha, \quad j = r+1, ..., p-1, \tag{4}$$

$$\Pr(\widehat{r} = r) \to 1, \quad r = p-1, \tag{5}$$

$$\Pr(\widehat{r} = r) \geq \theta_n \to 1 - (p-1-r)\,\alpha, \quad r < p-1. \tag{6}$$

Remark 2. Results in Theorem 3 are identical to those corresponding to an alternative (infeasible) estimator of r which bases every step of the procedure on true common trends. The reason is that, in every step, our method leads to valid common trends with probability tending to one.

Remark 3. Results in Theorem 3 are comparable to those of Theorem 12.3 of Johansen (1995) derived for standard cointegration, although there are a couple of differences. First, (5) is better than Johansen's result (who obtained the limit $1-\alpha$). However, for $r < p-1$ we obtained for $\Pr(\widehat{r} = r)$ a smaller lower bound than that achieved by Johansen $(1-\alpha)$. Nevertheless, note that the bound in (6) might not be strict and fits naturally with the upper bound given in (4).

Remark 4. Our procedure leads to straightforward estimation of the cointegrating space. In particular, suppose that the procedure is finalized in step k, $k = 1, ..., p-1$, so $\widehat{r} = p-k$. This step determines that the variables in $z_{(b)t} = (z_{pt}, z_{p-1,t}, ..., z_{p-k+1,t})'$ are possible common trends (valid ones with probability approaching one). Collect the rest of the elements of z_t in $z_{(a)t}$. Theorem 1 ensures that if $r = p-k$ and the elements of $z_{(b)t}$ are common trends, any set of $k+1$ variables formed by any of the variables in $z_{(a)t}$ and all those in $z_{(b)t}$ is always cointegrated. Thus there exists a $r \times k$ matrix B such that the components of $z_{(a)t} - Bz_{(b)t}$ have integration orders smaller than δ_p. B can be estimated by standard methods like ordinary least squares (OLS) or narrow band least squares (NBLS), the components in $z_{(a)t}$ being dependent variables and $z_{(b)t}$ the vector of regressors. NBLS provides consistent estimators in all situations.

Remark 5. Our procedure can be used to infer the dimension of possible cointegrating subspaces. In fact, it can be shown that the dimension of a cointegrating subspace can be inferred by applying our method to a set of cointegrating errors. The obvious difficulty is that those errors are unknown, although, they can be proxied by residuals. Then, the crucial issue is to justify that the inferential procedures employed ($\widehat{d}_i$ and $\widehat{\tau}_{j_1,\ldots,j_k}$) retain their properties when applied to residuals. A proper justification of this depends on the precise inference methods employed and goes beyond the scope of the present paper, although results in Hualde and Robinson (2006) indicate that this might be the case in many circumstances. We illustrate in Section 4 the problem of inferring the dimension of a cointegrating subspace and its estimation by means of a simple empirical example.

3. MONTE CARLO EVIDENCE

We investigate the finite sample performance of $\widehat{r}$ by means of a Monte Carlo experiment. Our analysis is based on 5000 replications of three series z_{it}, $i = 1, 2, 3$, of lengths $n = 256, 512$, generated according to five different DGP's, the numbering indicating the cointegrating rank (in all cases the u_t below are independent trivariate normal vectors such that $Var(u_{it}) = 1$, $i = 1, 2, 3$, $Cov(u_{it}, u_{jt}) = 0.5$, $i \neq j$): 0) $z_{it} = \Delta^{-.35} u_{it}$, $i = 1, 2, 3$; 1a) $z_{it} = \Delta^{-.35} u_{it}$, $i = 1, 2$, $z_{3t} = u_{3t}$; 1b) $z_{it} = \Delta^{-.35} u_{it}$, $i = 1, 2$, $z_{3t} + z_{2t} - z_{1t} = u_{3t}$; 2a) $z_{1t} = \Delta^{-.35} u_{1t}$, $z_{it} = u_{it}$, $i = 2, 3$; 2b) $z_{1t} = \Delta^{-.35} u_{1t}$, $z_{2t} - z_{1t} = u_{2t}$, $z_{3t} - z_{2t} - z_{1t} = u_{3t}$. Note that the a) cases reflect the situation where the z_{it} have distinct integration orders, whereas under b), the z_{it} share the same integration order, but they are cointegrated. In all cases the maximum integration order of the observables is 0.35, whereas cointegrating errors are always weak dependent. The fractional processes were generated by the algorithm designed by Davies and Harte (1987). We implement our procedure by using local Whittle estimation of the orders (Robinson, 1995b) and we test for $H_{j_1,\ldots,j_k}$ by the $\widehat{\Upsilon}_m^*$ semiparametric test statistic of Hualde and Velasco (2008) (using sizes $\alpha = .10, .05, .01$), $z_{j_k t}$ taking the role of the y_t variable in Hualde and Velasco's notation (this statistic is designed to be applied to Type II fractional processes, but it can be shown to be also adequate for Type I). Given that our inferential procedures are semiparametric, to check sensibility to bandwidth choice, we give results for $m = 55, 80$ and $m = 100, 150$, for $n = 256, 512$, respectively. We set $h_n = \log(m^{10^{-13}})$, so effectively h_n is indistinguishable from zero, while satisfying (2). We present in Table 1 the proportion of replications leading to each $\widehat{r}$ in the five different scenarios. Given the very adverse situation we face (with very small cointegrating gaps), our procedure, which is favoured when the cointegration is not trivial, performs very satisfactorily, especially for $r = 2$ as (5) suggests. In few cases $\widehat{r} = 0$

is chosen too frequently, but more often $\widehat{r} > r$ occurs, because our test statistic is usually oversized. However results improve substantially as n and m increase (note that given our design, we approach a parametric procedure as $m \to n/2$).

4. EMPIRICAL EXAMPLE

We apply our procedure to a trivariate series of 381 monthly observations (from January 1980 through September 2011) on Dubai, Brent, and West Texas Intermediate (WTI) oil prices (in $ per barrel) collected from the IMF Primary Commodity Prices database. We analyse log prices and identify Dubai, Brent and WTI log prices with z_{1t}, z_{2t} and z_{3t}, respectively. Similar data (for a much shorter time span) was employed in the empirical analysis of Robinson and Yajima (2002).

First, we estimate the integration orders of z_{it} by local Whittle. Given that nominal price series are usually characterised by strong nonstationarity (being this also confirmed by a preliminary graphical analysis), we use in our estimation first differences of the series. Estimates of the orders (adding back one to the values obtained from the differenced series) for bandwidth choices $m = 75+2i$, $i = 0, 1, ..., 10$, are reported in Table 2. Unlike the estimates presented by Robinson and Yajima (2002) (whose values were around 0.5), ours are in all cases very close to one. We also computed results for smaller bandwidths ($m \geq 15$), and in all cases the estimates were larger than 0.75 (most of them were between 0.8 and 1). Any of our bandwidth choices led to the identification of the common trend as $c_1 = 1$. Based on this, we computed Hualde and Velasco's (2008) test statistics $\widehat{\tau}_{1,2}$ and $\widehat{\tau}_{1,3}$, which in all cases took values much larger than critical ones, $\widehat{\tau}_{1,2}$ being always larger than $\widehat{\tau}_{1,3}$ (we do not report these statistics as they are not very informative). Thus, we concluded $\widehat{r} = 2$, being this conclusion also supported by any choice of $m \geq 15$. The cointegrating space was estimated by NBLS (with $m = 85$, although results are almost invariant to m) as the span of vectors $\widehat{a}_1 = (-.967, 1, 0)'$, $\widehat{a}_2 = (-.905, 0, 1)'$.

In Table 2, we also report results concerning the identification of a possible cointegrating subspace. In particular, e_{1t}, e_{2t} refer to cointegrating errors from the WTI-Dubai and Brent-Dubai cointegrating relations, respectively (proxied by NBLS residuals, $\widehat{e}_{it}$, $i = 1, 2$). Based on these residuals, we estimate by local Whittle the memory of e_{1t}, e_{2t} (and present these estimates in Table 2), both relations leading to cointegrating gaps close to 0.5. Next, using $\widehat{e}_{2t}$ as common trend, we computed Hualde and Velasco's (2008) test statistic (reported in Table 2 as $\widehat{\tau}_{e_2,e_1}$) to asses whether those errors are cointegrated, and it appears to be clearly the case, so, for our choices of m, the data support the existence of a cointegrating subspace. Note that $\widehat{d}_{g_1}$ reported in Table 2 is the local Whittle estimate of the memory of this subspace (computed from residuals of the NBLS regression of $\widehat{e}_{1t}$ on $\widehat{e}_{2t}$), being smaller than the estimated memories of e_{1t} or e_{2t}. Our conclusion is robust to other choices of m such that $m \geq 75$, whereas results for smaller m indicate, in general, that cointegration between cointegrating errors is trivial. In any case, given that $e_{1t} = z_{3t} - \beta_{31} z_{1t}$, $e_{2t} = z_{2t} - \beta_{21} z_{1t}$, if e_{1t} and e_{2t} are cointegrated, there exists θ such that $z_{3t} - \beta_{31} z_{1t} - \theta \left(z_{2t} - \beta_{21} z_{1t} \right)$ has reduced order. Then this subspace can be easily estimated by NBLS, choosing z_{3t} as dependent variable and z_{1t}, z_{2t}, as regressors. Choosing $m = 85$, we perform this estimation, concluding that the estimated subspace is the span of $\widehat{b} = (.078, -1.017, 1)'$. This estimate looks sensible because $\widehat{b}$ is almost identical to $\widehat{a}_2 - 1.017\widehat{a}_1$, the very small estimated coefficient corresponding to z_{1t} (0.078) implying that the cointegration between Brent and WTI possibly leads to a smaller memory than the one between any other pair of observable series.

5. CONCLUSION

We have proposed an automatic method to infer the cointegrating rank which does not rely on previous knowledge of memory or cointegrating characteristics of the vector of observables. The procedure can be applied irrespective of the cointegrating structure of the observables, it performs well in finite samples and it leads to straightforward estimation of the cointegrating space. Our method can be also employed to estimate possible cointegrating subspaces, and we illustrated this possibility by means of a simple empirical example.

REFERENCES

Chen, W. & C.M. Hurvich (2006) Semiparametric estimation of fractional cointegrating subspaces. *Annals of Statistics* 27, 2939-2979.

Davies, R.B. & D.S. Harte (1987) Test for Hurst effect. *Biometrika* 74, 95-101.

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

Flores, R. & A. Szafarz (1996) An enlarged definition of cointegration. *Economics Letters* 50, 193-195.

Gil-Alana, L.A. & J. Hualde (2009) Fractional integration and cointegration: an overview and an empirical application. *Palgrave Handbook of Econometrics*, Vol. 2, Applied Econometrics (K. Patterson and T.C. Mills, eds.). Palgrave, MacMillan, pp. 434-469.

Gomez-Biscarri, J. & J. Hualde (2011) Regression-based identification and estimation of cointegration structures. Preprint, Universidad Pública de Navarra.

Hualde, J. & P.M. Robinson (2006) Semiparametric estimation of fractional cointegration. STICERD Discussion paper EM/2006/502.

Hualde, J. & C. Velasco (2008) Distribution-free tests of fractional cointegration. *Econometric Theory* 24, 216-255.

Johansen, S. (1995) *Likelihood based inference in cointegrated vector autoregressive models*. Oxford University Press.

Kim, S.C. & P.C.B. Phillips (2000) Fully modified estimation of fractional cointegration models. Preprint, Yale University.

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

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

Robinson, P.M. (1995b) Gaussian semiparametric estimation of long-range dependence. *Annals of Statistics* 23, 1630-1661.

Robinson, P.M. (2008) Diagnostic testing for cointegration. *Journal of Econometrics* 143, 206-225.

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

APPENDIX

Proof of Theorem 2. The result follows by induction on showing: (i) $\Pr(c_1 = i_1) \to 1$ as $n \to \infty$; (ii) If for $r < p-1$, $k = 2, ..., p-r$, $\Pr\left(\cap_{l=1}^{k-1}\{c_l = i_l\}\right) \to 1$ as $n \to \infty$, then $\Pr\left(\cap_{l=1}^{k}\{c_l = i_l\}\right) \to 1$ as $n \to \infty$. We show (i) first. If $i_1 = p$

$$\Pr(c_1 = i_1) = \Pr\left(\bigcap_{i=1}^{p-1}\left\{g_{1n}(\widehat{d}_p - \widehat{d}_i) > h_n\right\}\right) \geq \sum_{i=1}^{p-1}\Pr\left(g_{1n}(\widehat{d}_p - \widehat{d}_i) > h_n\right) - (p-2). \tag{7}$$

For any $i = 1, ..., p-1$

$$\Pr\left(g_{1n}(\widehat{d}_p - \widehat{d}_i) > h_n\right) = \Pr\left((\widehat{d}_p - \widehat{d}_i - (d_p - d_i) > g_{1n}^{-1}h_n + d_i - d_p\right),$$

so the result follows by (2) because, as $n \to \infty$, $g_{1n}^{-1}h_n \to 0$, $d_i - d_p < 0$ and $\widehat{d}_p - \widehat{d}_i - (d_p - d_i) \to_p 0$. If $i_1 < p$, $\Pr(c_1 = i_1)$ equals

$$\Pr\left(\bigcap_{k=1}^{p-i_1}\left\{\bigcup_{i=1}^{p-k}\left\{g_{1n}(\widehat{d}_{p-k+1} - \widehat{d}_i) \leq h_n\right\}\right\}, \bigcap_{i=1}^{i_1-1}\left\{g_{1n}(\widehat{d}_{i_1} - \widehat{d}_i) > h_n\right\}\right)$$
$$\geq \sum_{k=i_1+1}^{p}\Pr\left(g_{1n}(\widehat{d}_k - \widehat{d}_{i_1}) \leq h_n\right) + \Pr\left(\bigcap_{i=1}^{i_1-1}\left\{g_{1n}(\widehat{d}_{i_1} - \widehat{d}_i) > h_n\right\}\right) - (p - i_1).$$

For $k = i_1 + 1, ..., p$, by (2)

$$\Pr\left(g_{1n}(\widehat{d}_k - \widehat{d}_{i_1}) \leq h_n\right) = \Pr\left(g_{1n}(\widehat{d}_k - \widehat{d}_{i_1} - (d_k - d_{i_1})) \leq h_n + g_{1n}(d_{i_1} - d_k)\right) \to 1,$$

as $n \to \infty$, because $g_{1n}(\widehat{d}_k - \widehat{d}_{i_1} - (d_k - d_{i_1})) = O_p(1)$ and $g_{1n}(d_{i_1} - d_k) = 0$ if $d_{i_1} = d_k$ or $\to \infty$ if $d_{i_1} > d_k$. The proof of (i) is concluded by showing that

$$\Pr\left(\bigcap_{i=1}^{i_1-1}\left\{g_{1n}(\widehat{d}_{i_1} - \widehat{d}_i) > h_n\right\}\right) \to 1 \text{ as } n \to \infty,$$

which holds by almost identical arguments to those employed in the proof of (7).

Next we show (ii). First, for $i_k < p-k+1$, $\Pr\left(\cap_{l=1}^{k}\{c_l = i_l\}\right)$ equals

$$\Pr\left(\bigcap_{l=1}^{k-1}\{c_l = i_l\}, \bigcap_{j=0}^{p-k+i_k}\left\{\bigcup_{i=1}^{p-k-j}\{\widehat{\tau}_{p,\dots,p-k+2,i} - \widehat{\tau}_{p,\dots,p-k+2,p-k-j+1} \le h_n\}\right\},\right.$$
$$\left.\bigcap_{i=1}^{i_k-1}\{\widehat{\tau}_{p,\dots,p-k+2,i} - \widehat{\tau}_{p,\dots,p-k+2,i_k} > h_n\}\right)$$

$$\begin{aligned}\ge\ & \Pr\left(\bigcap_{l=1}^{k-1}\{c_l = i_l\}\right) + \sum_{j=0}^{p-k+i_k}\Pr\left(\widehat{\tau}_{p,\dots,p-k+2,i_k} - \widehat{\tau}_{p,\dots,p-k+2,p-k-j+1} \le h_n\right) \\ & + \Pr\left(\bigcap_{i=1}^{i_k-1}\{\widehat{\tau}_{p,\dots,p-k+2,i} - \widehat{\tau}_{p,\dots,p-k+2,i_k} > h_n\}\right) - (p-k-i_k+2)\,.\end{aligned}$$

First, for $l = p-k+1, p-k, \dots, i_k+1$, $\Pr\left(\widehat{\tau}_{p,\dots,p-k+2,i_k} - \widehat{\tau}_{p,\dots,p-k+2,l} \le h_n\right)$ equals

$$\Pr\left(\widehat{\tau}_{p,\dots,p-k+2,i_k} \le h_n + \widehat{\tau}_{p,\dots,p-k+2,l}\right) \to 1 \text{ as } n \to \infty,$$

because $\widehat{\tau}_{p,\dots,p-k+2,i_k} = O_p(1)$ and $\widehat{\tau}_{p,\dots,p-k+2,l} = O_p(1)$ or $\to \infty$. Next

$$\begin{aligned} & \Pr\left(\bigcap_{i=1}^{i_k-1}\{\widehat{\tau}_{p,\dots,p-k+2,i} - \widehat{\tau}_{p,\dots,p-k+2,i_k} > h_n\}\right) \\ \ge\ & \sum_{i=1}^{i_k-1}\Pr\left(\widehat{\tau}_{p,\dots,p-k+2,i} - \widehat{\tau}_{p,\dots,p-k+2,i_k} > h_n\right) - (i_k-2)\,.\end{aligned}$$

For $i = 1, \dots, i_k - 1$, $\Pr\left(\widehat{\tau}_{p,\dots,p-k+2,i} - \widehat{\tau}_{p,\dots,p-k+2,i_k} > h_n\right)$ equals

$$\Pr\left(\widehat{\tau}_{p,\dots,p-k+2,i_k} < \widehat{\tau}_{p,\dots,p-k+2,i} - h_n\right) \to 1 \text{ as } n \to \infty,$$

by (2), because $\widehat{\tau}_{p,\dots,p-k+2,i_k} = O_p(1)$ and $\widehat{\tau}_{p,\dots,p-k+2,i}$ diverges to ∞ at a higher rate than h_n, to conclude the proof for $i_k < p-k+1$. Finally, for $i_k = p-k+1$, the proof is almost identical but simpler, so we omit it. Hence (ii) holds, to conclude the proof of the theorem.

Proof of Theorem 3. We show first (5). By the law of total probabilities

$$\Pr\left(\widehat{r}=p-1\right)=\sum_{i=1,i\neq i_1}^{p}\Pr\left(\widehat{r}=p-1,c_1=i\right)+\Pr\left(\widehat{r}=p-1,c_1=i_1\right). \qquad (8)$$

Given that for $i\neq i_1$, $\Pr\left(c_1=i\right)=o\left(1\right)$, the first term on the right hand side of (8) is $o\left(1\right)$. For any null hypothesis H_0, denote RH_0, AH_0, if H_0 is rejected or not rejected, respectively. Next, the second term on the right side of (8) equals

$$\Pr\left(\bigcap_{i=1}^{p-1}RH_{p,i},c_1=i_1\right)\geq\sum_{i=1}^{p-1}\Pr\left(RH_{p,i}\right)+\Pr\left(c_1=i_1\right)-\left(p-1\right),$$

so we conclude by Theorem 2, noting that $\Pr\left(RH_{p,i}\right)\rightarrow 1$ by consistency of $\widehat{\tau}_{p,i}$.

Next we show (4). Using again the law of total probabilities, by similar arguments to the ones above

$$\Pr\left(\widehat{r}=j\right)=\Pr\left(\widehat{r}=j,c_1=i_1,...,c_{p-j}=i_{p-j}\right)+o_p\left(1\right). \qquad (9)$$

First, for $j<p-1$, the first term on the right of (9) equals

$$\begin{aligned}&\Pr\left(\bigcap_{k=1}^{p-j-1}\left\{\bigcup_{i=1}^{p-k}AH_{p,...,p-k+1,i}\right\},\bigcap_{i=1}^{j}RH_{p,p-1...,j+1,i},c_1=i_1,...,c_{p-j}=i_{p-j}\right)\\ \leq\ &\Pr\left(\bigcap_{i=1}^{j}RH_{p,p-1...,j+1,i}\right)\leq\Pr\left(RH_{p,p-1...,j+1,l}\right),\end{aligned}$$

where $l\in\{1,...,j\}$ is such that $z_{pt},z_{p-1,t},...,z_{j+1,t},z_{lt}$, are not cointegrated, hence (4) for $j<p-1$ holds. For $j=p-1$ the proof is almost identical but simpler, so we omit it, to conclude (4).

Next we show (6). By previous arguments

$$\Pr\left(\widehat{r}=r\right)=\Pr\left(\widehat{r}=r,c_1=i_1,...,c_{p-r}=i_{p-r}\right)+o_p\left(1\right). \qquad (10)$$

The first term on the right of (10) equals

$$
\begin{aligned}
& \Pr\left(\bigcap_{k=1}^{p-r-1}\left\{\bigcup_{i=1}^{p-k} AH_{p,...,p-k+1,i}\right\}, \bigcap_{i=1}^{r} RH_{p,p-1...,r+1,i}, c_1 = i_1, ..., c_{p-r} = i_{p-r}\right) \\
\geq\; & \sum_{k=1}^{p-r-1} \Pr\left(\bigcup_{i=1}^{p-k} AH_{p,...,p-k+1,i}\right) + \Pr\left(\bigcap_{i=1}^{r} RH_{p,p-1...,r+1,i}\right) \qquad (11) \\
& + \Pr\left(c_1 = i_1, ..., c_{p-r} = i_{p-r}\right) - (p-r). \qquad (12)
\end{aligned}
$$

Clearly, $\Pr\left(\cup_{i=1}^{p-1} AH_{p,i}\right) \geq \Pr\left(AH_{p,l}\right) \to 1-\alpha$ as $n \to \infty$, where $l \in \{1, ..., p-1\}$ is such that z_{pt} and z_{lt} are not cointegrated. Similarly, any of the first $(p-r-1)$ terms on the right side of (11) can be bounded below by a term tending to $1-\alpha$, so (6) holds because $\Pr\left(\cap_{i=1}^{r} RH_{p,p-1...,r+1,i}\right) + \Pr\left(c_1 = i_1, ..., c_{p-r} = i_{p-r}\right) \to 2$ as $n \to \infty$, by Theorem 2 and consistency of the test.

Finally, we show (3). By previous arguments, for any $j < r$, $\Pr\left(\widehat{r} = j\right) \leq \Pr\left(\cup_{i=1}^{r} AH_{p,p-1...,r+1,i}\right) + o_p(1)$, so (3) holds because $\Pr\left(\cup_{i=1}^{r} AH_{p,p-1...,r+1,i}\right) \to 0$.

Table 1. Estimated ranks

	$\widehat{r}$	2	2	2	2	1	1	1	1	0	0	0	0
	n	256	256	512	512	256	256	512	512	256	256	512	512
r	$_{\alpha}\backslash^{m}$	55	80	100	150	55	80	100	150	55	80	100	150
	.10	.971	.974	1	1	.018	.015	.000	.000	.011	.011	.000	.000
$2a$	.05	.957	.951	.999	.999	.026	.024	.001	.001	.017	.025	.000	.000
	.01	.914	.882	.997	.995	.051	.060	.002	.003	.035	.058	.001	.002
	.10	.997	1	1	1	.002	.000	.000	.000	.001	.000	.000	.000
$2b$	.05	.996	.999	1	1	.002	.001	.000	.000	.002	.000	.000	.000
	.01	.994	.997	1	1	.003	.002	.000	.000	.003	.001	.000	.000
	.10	.254	.135	.236	.113	.720	.828	.760	.885	.026	.037	.004	.002
$1a$	.05	.188	.085	.172	.067	.772	.855	.822	.929	.040	.060	.006	.004
	.01	.111	.038	.090	.023	.800	.831	.896	.962	.089	.131	.014	.015
	.10	.168	.083	.147	.068	.802	.902	.846	.929	.030	.015	.007	.003
$1b$	.05	.126	.052	.106	.036	.829	.921	.881	.960	.045	.027	.013	.004
	.01	.078	.024	.054	.011	.846	.929	.919	.978	.076	.047	.027	.011
	.10	.111	.042	.106	.030	.288	.152	.242	.120	.601	.806	.652	.841
0	.05	.074	.023	.064	.015	.253	.107	.200	.073	.673	.870	.736	.905
	.01	.030	.005	.024	.003	.202	.063	.130	.029	.768	.932	.846	.963

Table 2. Estimated integration orders and test statistics

m	75	77	79	81	83	85	87	89	91	93	95
$\widehat{d}_1$	.978	1.00	1.02	1.01	1.03	1.03	1.05	1.05	1.07	1.09	1.08
$\widehat{d}_2$	.973	.994	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.06
$\widehat{d}_3$	.974	.990	1.00	1.01	1.02	1.02	1.03	1.03	1.05	1.05	1.05
$\widehat{d}_{e_1}$	.470	.468	.478	.491	.504	.496	.510	.521	.536	.547	.549
$\widehat{d}_{e_2}$	.575	.581	.586	.599	.588	.578	.587	.595	.605	.612	.603
$\widehat{\tau}_{e_2,e_1}$	354	448	314	263	314	238	197	130	84.5	52.0	28.8
$\widehat{d}_{g_1}$	.345	.339	.351	.359	.358	.359	.367	.378	.391	.403	.410