ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

ECONOMIA

ciclo XXVII

Settore Concorsuale di afferenza: 13/A5 Econometria

Settore Scientifico disciplinare: SECS/P05 Econometria

Essays in Macroecometrics:

methodological aspects and empirical applications

Vanessa Gunnella

Coordinatore del Dottorato
Matteo Cervellati

Relatori:
Roberto Golinelli
Luca Fanelli

Esame finale anno 2015

A *SUR-bounds* Panel Cointegration Test in the Presence of Cross-Section Dependence

Vanessa Gunnella[1]

Abstract

This paper introduces a new panel cointegration test. It extends Pesaran et al. (2001) bounds test by considering the individual regressions in a Seemingly Unrelated Regression (*SUR*) system. The algorithm to implement the test is developed and Monte Carlo simulation is used to analyze the properties of the test. The small sample properties of the test are remarkable, compared to its single equation counterpart. Size distortion is almost absent and power increases substantially. The use of the test is illustrated through a test of Purchasing Power Parity in a panel of EU15 countries.

JEL C12, C15, C23, C33
Keywords Panel cointegration; seemingly unrelated regression; bounds test; $I(1)$ and $I(0)$ regressors; Monte Carlo simulation.

---

[1]University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Email: vanessa.gunnella2@unibo.it
I wish to thank Jörg Breitung, Stefano Fachin, Luca Fanelli, Roberto Golinelli, Hashem Pesaran, Yongcheol Shin and all the participants of the 2^{nd} Rimini Time Series Workshop, the DIW Macroeconometric Workshop and the BGSE Macro/Econometrics/Finance Workshop for useful comments and suggestions.

1 Introduction

In time series analysis, cointegration is a statistical property that is used to detect meaningful relationships between non-stationary variables and thus, to test for the existence of long-run links.

Panel cointegration tests have been introduced in the literature in order to improve the power of uni-equational tests. The extension to a cross-section of N units - countries, regions, firms, etc - increases the number of observations and allows to exploit more information. Also, the panel setting makes it possible to test for hypotheses that should hold for a group of units.
Even though panel cointegration analysis seems to be overlooked in recent times, in the last years many papers have used this technique in order to address empirical questions in a variety of economic fields (see for example Chong et al., 2012, Herzer et al., 2012 and Rassenfosse and Potterie, 2012).

Nevertheless, the approach is affected by two main issues: how to combine the information coming from all the units of the panel and how to deal with the cross-sectional dependence.

Regarding the first issue, many testing methodologies proposed in the literature have opted either for estimating a single test's parameter pooling the observations (e.g. Kao, 1999) or for estimating separate equations - one for each individual - and pooling the test's statistics (see Pedroni, 2004; Westerlund, 2007 and Larsson et al., 2001 among the others).
However, neglecting the parameters' heterogeneity can lead to biased estimates, whereas producing unique statistics do not give information about how many and which units of the panel are cointegrated.

On the other hand, assuming cross-section independence leads to inefficiency and size distortion of the tests (see O'Connell, 1998).
In many recent contributions, cross-section dependence is modeled with a common factors representation (Gengenbach et al., 2006; Westerlund and Edgerton, 2008; Gengenbach et al., 2008). Nevertheless, the small cross-sectional dimension of the panel can compromise the estimation of the common factors.

The test proposed here, called the *SUR-bounds* test, is conceived to tackle the above mentioned problems.
Heterogeneous parameters for the N testing regressions are estimated in a *SUR* (*Seemingly Unrelated Regressions*) system with *FGLS*. This allows to take into account unobserved common factors that contemporaneously affect all the units of the panel providing, at the same time, unit-specific test statistics. Moreover, the approach is particularly suited when the number of individuals of the panel is small relatively to the number of time series

observations.[2]

FGLS estimation of a *SUR* system has been adopted in the panel literature in the context of unit root testing (Breuer et al., 2002; Breitung and Das, 2005), in the cointegration rank test proposed by Breitung (2005) and in the estimation of cointegrating regressions (Mark et al., 2005).
In this paper the methodology is applied to extend Pesaran et al. (2001)'s (PSS hereafter) single equation bounds test. This test verifies the presence of a long-run relationship by testing for the joint significance of the lagged variables in an *ECM* representation. The most appealing feature of the test is that, differently from other tests, it allows the regressors to be $I(1)$, $I(0)$ or mutually cointegrated.
The *SUR* extension allows to exploit the information coming from the variance covariance matrix and hence it leads to increased power properties.

The distribution of the new test statistics is neither pivotal, nor boundedly pivotal because it depends on the covariance matrix of the panel. Thus critical values have to be simulated for each dataset under investigation by stochastic simulation. Hence, an algorithm for critical values' simulation is provided in order to allow the practitioner to apply the test.[3] The properties of the new panel test are compared with those of PSS. The results of the Monte Carlo simulation are unequivocally in favor of the new test. Even though both tests are found to be not dramatically distorted - although with better size properties of the new test in high dependence scenarios - power gains of the *SUR-bounds* test are widespread, especially in the presence of high cross-sectional correlation. A Monte Carlo experiment shows that the *SUR-bounds* test has superior power in comparison with Westerlund (2007) panel cointegration test.
The use of the test is shown in an empirical application of the test of the *PPP*-hypothesis in EU15 countries.
The rest of the paper proceeds as follow. Section 2 outlines the methodology: PSS bounds test is presented, then the panel extension is introduced and the testing procedure is illustrated.
Section 3 describes the algorithm to simulate the critical values whereas in Section 4 Monte Carlo experiments are performed in order to investigate size and power properties, comparing the new test with PSS single equation test and Westerlund (2007) panel test, under various scenarios. Section 5 provides the empirical application and conclusions are drawn in Section 6.

[2] In order to consistently estimate the elements of the residuals' variance-covariance matrix, the time dimension should be significantly larger than the size of the panel, i.e. $T >> N$.

[3] The development of the command is now object of research. The idea is to make such algorithm available to the scientific community.

2 Single equation test and its extension to the panel data framework

The proposed test is an extension to the panel setting of the bounds test of PSS. In this testing procedure, cointegration is verified considering the joint significance of the lagged variables of an unrestricted conditional *ECM*.

The panel test is characterized by the same features of the single equation counterpart.

On the one hand, it assumes the existence of only one cointegrating relation and weakly exogenous regressors.
The first restriction characterizes all the single equation and panel tests that abstract from a *VAR-VECM* approach to cointegration *à la* Johansen (1991)[4].
The second one is required for the *ECM*-based tests. If the regressors are not weakly exogenous, the conditional model do not contain the necessary information to test for cointegration (see Zivot (2000)).
However, as Pesaran and Shin (1998) show, a potential endogeneity problem can be circumvented adding enough lags of the regressors.

On the other hand, the test shows two interesting features: the *a priori* knowledge of the order of integration of the regressors is not needed and the regressors are allowed to be mutually cointegrated. In fact, two sets of critical values are computed: one corresponding to the cases in which all the regressors are $I(0)$ and the other considering the possibility that all the regressors are $I(1)$. This critical values provide respectively the lower and the upper bound of the test.

The extension to the panel setting of this test is carried out by estimating a *SUR* system featuring N individual *ECM*s.
This allows to estimate heterogeneous, individual specific parameters for each equation and therefore to test separate null hypotheses.
At the same time, the information provided by the cross-equation error covariance structure is exploited leading to more efficient estimators and consequently more powerful test statistics, conditional on the use of "correct" critical values.

[4] In the literature of panel cointegration tests, see for instance Pedroni (2004) and Westerlund (2007)

2.1 The individual bounds test

The *VECM* representation of the $VAR(p)$ model underlying the data generating process of the $(k+1)$-vector of variables $\{z_t\}_{t=1}^{\infty}$ is the following:

$$\Delta \mathbf{z}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{\Pi} \mathbf{z}_{t-1} + \sum_{j=1}^{p-1} \mathbf{\Gamma}_j \Delta \mathbf{z}_{t-j} + \boldsymbol{\varepsilon}_t \tag{1}$$

where $\mathbf{a}_0$ and $\mathbf{a}_1$ are $(k+1)$-vectors of intercept and trend coefficients, $\mathbf{\Pi}$ is the long-run multipliers matrix and $\mathbf{\Gamma}_j$ is the short run coefficient matrix of the differenced variables with lag j.

Under assumption 1-5 of PSS[5], the bounds test is applied to the conditional model $E\Big[y_t|\mathbf{x}_t, \{\mathbf{z}_{t-j}\}_{j=1}^{t-1}\Big]$ where $\mathbf{z}_t = (y_t, \mathbf{x}_t)'$. Hence, the conditional *ECM* is:

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx,x} \mathbf{x}_{t-1} + \sum_{j=1}^{p-1} \boldsymbol{\psi}_j' \Delta \mathbf{z}_{t-j} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + u_t \tag{2}$$

To test for no cointegration, the joint null hypothesis is:

$$H_0^{\pi_{yy}} : \pi_{yy} = 0 \qquad \wedge \qquad H_0^{\boldsymbol{\pi_{yx,x}}} : \boldsymbol{\pi_{yx,x}} = \mathbf{0}' \tag{3}$$

whereas the alternative hypothesis is:

$$H_1^{\pi_{yy}} : \pi_{yy} \neq 0 \qquad \vee \qquad H_1^{\boldsymbol{\pi_{yx,x}}} : \boldsymbol{\pi_{yx,x}} \neq \mathbf{0}' \tag{4}$$

This is implemented with the usual Wald test or the F-test for the joint significance of the coefficients of the lagged variables $\mathbf{z}_{t-1}$. Moreover, PSS propose a bounds procedure also for the test proposed by Banerjee et al. (1998) which is based on testing the significance of the coefficient of the lagged dependent variable with a t-test.
However, the Wald-statistic and the F-statistic have asymptotic null distributions that depend on the deterministic specification of (2).[6]
The limiting distributions of the statistics depend also on whether the variables within the vector $[\mathbf{x}_t]_{t=1}^{\infty}$ are cointegrated and on the number of the regressors (k). Two polar cases are studied - namely, $\{\mathbf{x}_t\} \sim I(1)$ when the

[5] See mentioned paper.

[6] See Pesaran et al. (2001) for the five different specifications and for the derivation of the limiting distribution of the Wald statistic.

rank is equal to k and $\{\mathbf{x}_t\} \sim I(0)$ when the rank is equal to 0 - and the critical values of the statistics are derived through stochastic simulation.
These two sets of critical values represent, respectively, the upper and the lower bounds for the test. That is, if the statistic is higher than the upper bounds then $\{y_t\}_{t=1}^{\infty}$ and $\{\mathbf{x}_t\}_{t=1}^{\infty}$ are cointegrated.
On the other hand, if the value of the statistic is smaller than the lower bounds the processes are not cointegrated.
Finally, if the statistic falls within the bounds, the $\{\mathbf{x}_t\}_{t=1}^{\infty}$ are mutually cointegrated and an analysis of their orders of integration is needed to reach a meaningful conclusion.

2.2 The panel model and the panel test

In this section, the extension of the bounds test to the panel environment is illustrated.

In order to take into account the heterogeneity of the parameters and, at the same time, to consider the cross-sectional dependence across the variables involved, the individual testing regressions (2) for the N units of the panel are estimated in a *Seemingly Unrelated Regressions* (*SUR*), as Breuer et al. (2002) do in the context of unit root testing.
This method treats the equations separately, but, assuming the errors to be correlated across panel units, it estimates for each unit specific parameters applying the *FGLS* procedure.

Hence, a system of unrestricted conditional *ECM* is considered. For simplicity of notation, deterministic constants and trends are not included and the coefficients associated with y_{t-1} and $\mathbf{x}_{t-1}$ are renamed, respectively with π_y and π_x.

$$\begin{cases} \Delta y_{1t} = \pi_{1y} y_{1,t-1} + \boldsymbol{\pi}_{1x}\mathbf{x}_{1,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\psi}_{1j}' \Delta\mathbf{z}_{1,t-j} + \boldsymbol{\omega}_1' \Delta\mathbf{x}_{1t} + u_{1t} \\ \Delta y_{2t} = \pi_{2y} y_{2,t-1} + \boldsymbol{\pi}_{2x}\mathbf{x}_{2,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\psi}_{2j}' \Delta\mathbf{z}_{2,t-j} + \boldsymbol{\omega}_2' \Delta\mathbf{x}_{2t} + u_{2t} \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ \Delta y_{Nt} = \pi_{Ny} y_{N,t-1} + \boldsymbol{\pi}_{Nx}\mathbf{x}_{N,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\psi}_{Nj}' \Delta\mathbf{z}_{N,t-j} + \boldsymbol{\omega}_N' \Delta\mathbf{x}_{Nt} + u_{Nt} \end{cases} \tag{5}$$

The matrix form of the i^{th} individual equation is:

$$\Delta\mathbf{y}_i = \mathbf{Z}_{i,-1}\boldsymbol{\pi}_{i,yx}^{*} + \Delta\mathbf{Z}_{i-}\boldsymbol{\psi}_i + \mathbf{u}_i \qquad i = 1, \ldots, N \tag{6}$$

where $\Delta\mathbf{y}_i \equiv (\Delta y_{i1}, \ldots, \Delta y_{iT})'$, $\mathbf{Z}_{i,-1} \equiv (\mathbf{z}_{i0}, \ldots, \mathbf{z}_{i,T-1})'$, $\Delta\mathbf{X}_i \equiv (\Delta\mathbf{x}_{i1}, \ldots, \Delta\mathbf{x}_{iT})'$, $\Delta\mathbf{Z}_{i,-j} \equiv (\Delta\mathbf{z}_{i,1-j}, \ldots \Delta\mathbf{z}_{i,T-1})'$, $\Delta\mathbf{Z}_{i-} \equiv (\Delta\mathbf{X}_i, \Delta\mathbf{Z}_{i,-1}, \ldots, \Delta\mathbf{Z}_{i,1-p})$, $\boldsymbol{\psi}_i \equiv (\boldsymbol{\omega}_i', \boldsymbol{\psi}_{i1}, \ldots, \boldsymbol{\psi}_{i,p-1})'$, $\mathbf{u}_i \equiv (u_{i1}, \ldots, u_{iT})'$ and

$$\boldsymbol{\pi}^*_{i,yx} = \begin{pmatrix} \mathbf{0}' \\ \mathbf{I}_{k+1} \end{pmatrix} \begin{pmatrix} \pi_{iy} \\ \boldsymbol{\pi}'_{ix} \end{pmatrix} \tag{7}$$

Thus, considering the system (5) in a stacked form, the following equation is obtained:

$$\Delta \mathbf{y} = \mathbf{Z}_{-1}\boldsymbol{\pi}^*_{yx} + \Delta\mathbf{Z}_{-}\boldsymbol{\psi} + \mathbf{u} \tag{8}$$

where $\mathbf{y} \equiv (\mathbf{y}_1, \ldots, \mathbf{y}_N)$, $\boldsymbol{\pi}^*_{yx} \equiv (\boldsymbol{\pi}^*_{1,yx}, \ldots, \boldsymbol{\pi}^*_{N,yx})$, $\mathbf{u} \equiv (\mathbf{u}_1, \ldots, \mathbf{u}_N)$,

$$\mathbf{Z}_{-1} \equiv \begin{bmatrix} \mathbf{Z}_{1,-1} & 0 & \cdots & 0 \\ 0 & \mathbf{Z}_{2,-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{Z}_{N,-1} \end{bmatrix} \text{ and } \Delta\mathbf{Z}_{-} \equiv \begin{bmatrix} \Delta\mathbf{Z}_{1-} & 0 & \cdots & 0 \\ 0 & \Delta\mathbf{Z}_{2-} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta\mathbf{Z}_{N-} \end{bmatrix} \tag{9}$$

Assumption 1-5 of PSS hold and, furthermore, it is further assumed:

Assumption 1. There is no cross-unit cointegration among the stacked variables belonging to the matrix $\mathbf{Z}_{-1}$[7].

Moreover, the following assumptions on the stacked error vector $\mathbf{u}$ are imposed:

Assumption 2. $E[\mathbf{u}|\mathbf{Z}_{-1}, \mathbf{Z}_{-}] = \mathbf{0}$ and $E[\mathbf{u}\mathbf{u}'|\mathbf{Z}_{-1}, \mathbf{Z}_{-}] \equiv \boldsymbol{\Omega}$ is positive definite.

Assumption 3. $E[u_{it}u_{js}|\mathbf{Z}_{-1}, \mathbf{Z}_{-}] = \sigma_{ij}$ for $t = s$ and 0 otherwise.

It follows that for the *t-th* observation, the covariance matrix is:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN} \end{bmatrix} \tag{10}$$

so that $\boldsymbol{\Omega} = \Sigma \otimes \mathbf{I}_T$.

In a first step, the covariance matrix $\boldsymbol{\Omega}$ is estimated with LS and $\hat{\boldsymbol{\Omega}}$ is incorporated in the estimators of the parameters of interest $\hat{\boldsymbol{\pi}}^{*i}_{y,x}$ (i=1, ..., N) in the second step.[8]

[7]As Banerjee et al. (2004) point out, panel cointegration test can be oversized in the presence of cross-unit cointegration.

[8]Alternatively, the covariance matrix could be estimated with the procedure proposed by Mark et al. (2005)

Therefore, the following set of separate null and alternative hypothesis is tested with unit-specific Wald or F statistics:

$$\begin{array}{cc} H_0^{1\pi_{1y}} \wedge H_0^{1\boldsymbol{\pi}_{1x}} & H_1^{1\pi_{1y}} \vee H_1^{1\boldsymbol{\pi}_{1x}} \\ H_0^{2\pi_{2y}} \wedge H_0^{2\boldsymbol{\pi}_{2x}} & H_1^{2\pi_{2y}} \vee H_1^{2\boldsymbol{\pi}_{2x}} \\ \vdots & \vdots \\ H_0^{N\pi_{Ny}} \wedge H_0^{N\boldsymbol{\pi}_{Nx}} & H_1^{N\pi_{Ny}} \vee H_N^{N\boldsymbol{\pi}_{Nx}} \end{array} \tag{11}$$

with $H_0^{i\pi_{iy}} : \pi_{iy} = 0$ and $H_1^{i\pi_{iy}} : \pi_{iy} \neq 0$ and $H_0^{i\pi_{ix}} : \pi_{ix} = 0$ and $H_1^{i\pi_{ix}} : \pi_{ix} \neq 0$.

Whereas, the following null and alternative hypothesis are tested with unit-specific t statistic:

$$\begin{array}{cc} H_0^1 : \pi_{1y} = 0 & H_1^1 : \pi_{1y} \neq 0 \\ H_0^2 : \pi_{2y} = 0 & H_1^2 : \pi_{2y} \neq 0 \\ \vdots & \vdots \\ H_0^N : \pi_{Ny} = 0 & H_1^N : \pi_{Ny} \neq 0 \end{array} \tag{12}$$

Moreover, a joint significance test on the coefficients of the lagged variables of all the equations is proposed. That is, the null hypothesis of (11) are tested jointly with an F test and a Wald test (*sys*-F and *sys*-W test hereafter). Under the null, there is no cointegration in all the panel, whereas under the alternative at least one of the equation of the panel is cointegrated. These tests, as most of the panel tests (see Westerlund (2007), Pedroni (2004) among the others), give a unique, synthetic information on the panel as a whole. While this constitutes a limitation of the other panel tests, in the present context this represents an additional information.

Apart from depending on the number of regressors, on the deterministic specification, on the panel dimension and on the cointegration rank of $[\mathbf{x}_t^i]_{t=1}^{\infty}$ $(i = 1, \ldots, N)$, the null distributions of the test statistics are specific to the estimated covariance matrix $\hat{\boldsymbol{\Omega}}$.

That is why the critical values are series specific and have to be computed through stochastic simulation. This would not represent an obstacle to the implementation of the test by practitioners because a program (e.g. Stata .ado file) will be provided in order to make it possible to easily perform the test on other datasets and will thus spread the applicability of the test.

In the next session the algorithm to simulate critical values is illustrated.

3 Critical values' simulation

In what follows, the steps of the procedure put forth to produce the data-set specific critical values are illustrated. As in Breuer et al. (2002), Monte Carlo

methods are applied.
The simulated data are generated from the estimated covariance matrix and coefficients of the model under the null hypothesis. Therefore, starting from randomly generated errors and regressors series - for every equation of the system - dependent variables series are derived. Then, the test statistics are computed with the simulated sample. Their discrete distributions are obtained performing a suitable number of replications and the critical values correspond to the percentiles of interest of the distributions.

The procedure consists of the following steps[9]:

1. Choose the lag orders of the dependent variable and of the regressor[10]. with a lags selection method in each regression of the system.[11]

2. Estimate the SUR system of the testing regressions under the null hypothesis, i.e.:

$$\begin{cases} \Delta y_{1t} = \sum_{j=1}^{p_1-1} \psi_{1yj} \Delta y_{1,t-j} + \sum_{j=1}^{q_1-1} \psi_{1xj} \Delta x_{1,t-j} + \omega_1 \Delta x_{1t} + u_{1t} \\ \Delta y_{2t} = \sum_{j=1}^{p_2-1} \psi_{2yj} \Delta y_{2,t-j} + \sum_{j=1}^{q_2-1} \psi_{2xj} \Delta x_{2,t-j} + \omega_2 \Delta x_{2t} + u_{2t} \\ \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \Delta y_{Nt} = \sum_{j=1}^{p_N-1} \psi_{Nyj} \Delta y_{N,t-j} + \sum_{j=1}^{q_N-1} \psi_{Nxj} \Delta x_{N,t-j} + \omega_N \Delta x_{Nt} + u_{Nt} \end{cases} \tag{13}$$

   in order to get $\hat{\psi}_{iyj,H_0}$ $(j = 1, ..., p_i)$, $\hat{\psi}_{ixj,H_0}$ $(j = 1, ..., q_i)$ and $\hat{\omega}_i$ for $i = 1, .., N$ and the estimated covariance matrix of the errors of the system, $\hat{\Sigma}_{H_0}$.

3. For $i = 1, ..., N$, generate $u_{it}^{(s)}$ and $x_{it}^{(s)}$. The first, as random draw from standard normal distribution multiplied by the Cholesky decomposition of $\hat{\Sigma}_{H_0} = SS'$. The latter, generated from $x_{it}^{(s)} = x_{i,t-1}^{(s)} + \epsilon_{it}^{(s)}$, in the case that x_{it} is assumed to be purely $I(1)$ or from $x_{it}^{(s)} = \epsilon_{it}^{(s)}$ if x_{it} is purely $I(0)$.[12] $\epsilon_{it}^{(s)}$ is a draw from independent standard normal variables.

[9]For the sake of simplicity of notation, only one regressor is considered.

[10] As Pesaran et al. (2001) clarify, it is possible to have heterogeneous lag orders for the variables without compromising the asymptotic results of their test, so that the model is an $ARDL(p, q_1, q_2, ..., q_k)$.

[11]Pesaran and Shin (1998) demonstrate through Monte Carlo experiments that the estimators of the $ARDL$ model with lag structure selected with Schwartz Bayesian Criterion perform slightly better than those selected with Akaike Information Criterion.

[12]These two cases characterize, respectively, the upper and the lower critical values bounds, are described in Section 2.

4. Generate $y_{it}^{(s)}$ with the estimated parameters and the generated residuals and regressors, imposing the null, i.e.:

$$\Delta y_{it}^{(s)} = \sum_{j=1}^{p_i-1} \hat{\psi}_{iyj,H_0} \Delta y_{i,t-j}^{(s)} + \sum_{j=1}^{q_i-1} \hat{\psi}_{ixj,H_0} \Delta x_{i,t-j}^{(s)} + \hat{\omega}_i \Delta x_{it}^{(s)} + u_{it}^{(s)} \quad (14)$$

and

$$y_{it}^{(s)} = y_{i,t-1}^{(s)} + \Delta y_{it}^{(s)} \quad (15)$$

Notice that here the initial conditions for y_{it} and x_{it} can be set to zero generating extra observations and discarding the first ones.

5. Estimate π_{iy} and π_{ix} $(i = 1, ..., N)$ from the *SUR* system of unrestricted testing regressions using the simulated sample and the estimated covariance matrix, $\hat{\Sigma}_{H_0}$. The i^{th} element of the system is:

$$\Delta y_{it}^{(s)} = \pi_{iy} y_{i,t-1}^{(s)} + \pi_{ix} x_{i,t-1}^{(s)} + \sum_{j=1}^{p_i-1} \psi_{iyj} \Delta y_{i,t-j}^{(s)} + \sum_{j=1}^{q_i-1} \psi_{ixj} \Delta x_{i,t-j}^{(s)} + \omega_i \Delta x_{it}^{(s)} + u_{it} \quad (16)$$

6. Compute the Wald and the F-statistics corresponding to the joint null hypothesis $H_0^{yi} : \pi_{iy} = 0 \wedge H_0^{xi} : \pi_{ix} = 0$, the t-statistics for the null $H_0^{yi} : \pi_{iy} = 0$ and the sys-W and the sys-F statistics for the null of all zero coefficients on the lagged variables of the system.

Points 3-6 are repeated the necessary number of times in order to obtain discrete distributions of the test statistics and the critical values are taken from the corresponding percentile of interest.

4 Monte Carlo simulations

In this section, small sample size and power properties of the new test are evaluated and are compared to those of Pesaran et al. (2001)'s single equation test - of which this test represents the extension and to those of a popular *ECM*-based panel test (Westerlund, 2007).[13]
In order to confront the *SUR-bounds* test with the bounds test and with Westerlund's test, the mean size and power of the *SUR-bounds* test are considered, because the former deals with only one individual, whereas the latter provides synthetic panel statistics.

[13]Comparisons with single equation and panel residual-based tests are not considered, since they have been already performed, among the others, by Pesavento (2004) for the former case and by Westerlund (2007) for the latter case.

The DGP adopted borrows from that of Westerlund's, but it has been modified in order to address specific issues that are illustrated later on.
In this context, one non-stationary $I(1)$ regressor is taken into account and it is assumed that the order of integration of the regressor is known *a priori*.[14] Therefore, only the upper bound of the *SUR-bounds* test is considered in order to make decisions on the outcome of the test.

4.1 Baseline scenario

This simulation experiment compares the tests on the ground of cross-sectional dependence and assumes the particular structure underlying the *SUR-bounds* test, namely, the presence of cross-correlation among the error terms of the individual regressions.
The DGP is the following:

$$\Delta y_{it} = \pi_{i,yy} y_{i,t-1} + \pi_{i,yx} x_{i,t-1} + \omega_i \Delta x_{it} + u_{it} \tag{17}$$

with

$$x_{it} = x_{i,t-1} + \varepsilon_{it} \tag{18}$$

Here, no lags and deterministic component are included[15].
In this scenarios it is assumed that x_{it} is strictly exogenous, therefore $\omega_i = 0$. Hence, the errors of the N regressors are assumed to be uncorrelated with $u_{it}, \quad i = 1, .., N$ and uncorrelated with each other. On the other hand the errors of equation (17) are assumed to have as covariance matrix the unit diagonal $N \times N$ matrix Σ. Each off-diagonal element of Σ represents the errors' covariance between two units and it is drawn from a uniform distribution. As in Breuer et al. (2002), according to the degree of correlation, the support of the distribution is defined between three different ranges: $0.20 - 0.30$ for low cross-sectional dependence, $0.45 - 0.55$ for medium cross-sectional dependence and $0.70 - 0.80$ for high cross-sectional dependence.
Therefore, a covariance matrix for the errors of equation (17) and equation (18) is built such that:

$$\underbrace{\breve{\Sigma}}_{2N \times 2N} = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{I_N} \end{bmatrix} \tag{19}$$

and the residuals are generated as independent draws from a standard normal distribution multiplied by the Cholesky decomposition of the matrix $\breve{\Sigma}$.

[14]This implies that the degenerate case in which the coefficient of the dependent variable is equal to zero while the coefficient of the regressor is different from zero is excluded, as it will be illustrated in what follows.

[15]The deterministic specification corresponds to model I of PSS.

The parameters $\pi_{i,yy}$ and $\pi_{i,yx}$ are set to zero in order to study the size properties.
Under the alternative hypothesis and the fact that $x_{it} \sim I(1)$ $\pi_{i,yx}$ do not depend on any parameter of the marginal model for x_{it} so that it is unrestricted.
Hence, it is more suitable to consider an alternative representation of the model (17), that is:

$$\Delta y_{it} = \alpha_i(\beta_{iy} y_{i,t-1} + \beta_{ix} x_{i,t-1}) + u_{it} \tag{20}$$

where the parameters are set as follows: β_{iy} is set to one for all the units, $\beta_{ix} \sim N(0,1)$ and $\alpha_i = 0$ under the null and is drawn from $U(-0.01, -0.03)$ and $U(-0.07, -0.09)$ for two different case studies for the power analysis (one closer to the null, the other more distinct from the null).
Notice that, under the null, both parameters are zero whereas under the alternative the coefficient are both different from zero.

The testing regressions are estimated choosing the individual lag orders with the *SBC* criterion and the maximum number of lags varies with T, according to the rule suggested by Schwert (1989) $\left\lfloor 12 \times \left[(T/100)^{1/4} \right] \right\rfloor$.

Since critical values of the *SUR-bounds* test are specific to the covariance matrix of each correlation scenario, they have to be computed for the three cases of low, medium and high cross-correlation. These critical values are taken as reference for the decision of rejection of the null hypothesis in the $SUR-bounds$ test.
The simulations are performed considering several cross-section and time dimensions that is $N = 5, 10, 20$ and $T = 50, 100, 200$. Therefore, 27 different environments are analyzed (3 types of cross-sectional dependence $\times$ 3 cross-section dimensions $\times$ 3 time dimensions). Furthermore, PSS bounds test's critical values have to be computed for $T = 50, 100, 200$ since the critical values of their paper refer to a sample size of 1000 observations[16].

The 0.05 lower and upper critical values for each experiment are shown in Table 2. As it is possible to see, the magnitude of the critical values decreases with the time dimension of the panel. On the contrary, the critical values increase as N is growing. Moreover, the critical values are bigger the higher is the cross-sectional dependence between the units of the panel.
Compared to PSS critical values presented in Table 1, the critical values are substantially higher, even for low cross-correlation.

Table 3 reports size results of *SUR-bounds* and PSS bounds tests for a nominal size equal to 0.05. For $\alpha_i = 0$, size magnitudes are reported. Both

[16]Narayan (2005) provides critical values for $T = 30 - 80$, but only for case II to V of PSS, that is excluding the "no intercept, no trend" case considered here.

tests show correct size or at the most minimal distortion. PSS test seems to be slightly more correct for low cross-correlation scenarios, whereas the *SUR-bounds* has better size properties for cross-correlation range 0.45-0.55 and 0.70-0.80.
On the other hand, the *sys-F* statistics overperforms the *sys-W*, except for the high cross-correlation case.

The power results are shown in Table 4-5. The first table reports respectively the power and the size-corrected power of the tests comparing the *SUR-bounds* tests with the PSS tests for the case of close to the null alternative, whereas the latter assumes the 0.07-0.09 alternative.
The comparison is totally in favor of *SUR-bounds* test. The power of the new test is higher than that PSS test in every scenario. Although both tests perform poorly when $\alpha = U(-0.01, -0.03)$ (the highest power is reached by the *SUR-bounds* tests under the high correlation, N=10, T=200 scenario), *SUR-bounds* test shows good power properties, in particular in the case of high cross-sectional dependence and for $T = 100, 200$.
Particularly, the panel tests seem to perform best, reaching the desired power in the high cross-sectional correlation scenarios.

4.2 Serial correlation

The second scenario is designed to analyze the effect of serial correlation in the error terms of the testing regressions. Hence, the N error processes are modeled as follow:

$$u_{it} = e_{it} + \rho_{1i}e_{i,t-1} + \rho_{2i}e_{i,t-2} \tag{21}$$

where the autoregressive parameter ρ_{ji} $j = 1, 2$ are heterogeneous across units and are uniformly distributed between 0.4 and 0.5 in case of positive serial correlation and between -0.4 and -0.5 in case of negative serial correlation.
Size magnitudes are shown in Table 6. In the $MA(1)$ case, while the sizes of the individual tests are almost preserved, the joint tests suffer from some distortion, especially the *sys-W* test.
Considering a higher serial correlation ($MA(2)$), the size distortion increases, as expected, but not so dramatically.

4.3 Mixed Panel

This scenario is meant to verify the power properties in the presence of mixed alternative, i.e. when under the alternative hypothesis some number of units (n) of the panel are simulated under no-cointegration and the remaining under cointegration $(N - n)$:

$$\Delta y_{it} = u_{it} \quad \text{for } i = 1, ..., n$$

and

$$\Delta y_{it} = \alpha_i(y_{i,t-1} + \beta_{ix} x_{i,t-1}) + u_{it} \quad \text{for } i = n+1, ..., N$$

For this experiment $\alpha_i \sim U(-0.09, -0.07)$, $N = 10$ and $T = 100$. Table 7, reports power magnitudes for various sizes of n. As it is possible to see, individual tests suffer from negligible loss of power, increasing with n. Instead, for the panel statistics, the power decrease is more pronounced, but they can nevertheless preserve nominal power in M and H scenarios, as long as $n \leq (N/2)$.

4.4 Comparison with Westerlund (2007)

In this subsection the *SUR-bounds* test is compared with Westerlund (2007) panel cointegration test. It is an ECM-based test which treats cross-sectional dependence with a bootstrap procedure as in Chang (2004). The author formulated four test statistics: two group statistics G_τ and G_α and two panel statistics P_τ and P_α (see mentioned paper for details).
Data have been simulated under the baseline scenario described above, considering a panel dimension of $N = 10$ and $T = 100$. Results are reported in Table 8.
For what regards the sizes, Westerlund's τ statistics are quite distorted, whereas the α statistics are undersized.
Comparing the power properties of the two tests, the individual statistics of the *SUR-bounds* are those performing worst. However, the *SUR* panel statistics beat the Westerlund's statistics almost in all cases, but especially when the cross-sectional dependence is substantial (M and H).

5 Empirical application

As illustrative example, the new methodology is applied to the PPP hypothesis testing in a panel of EU15 countries[17]for the quarterly sample 1974q1-1998q3.
The example is a classic empirical application in the unit-root and cointegration hypothesis testing literature (see Breuer et al. (2002) and Pedroni (2004) for what regards the panel literature) and it has been chosen for its

[17]EU15 comprises: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom.

popularity.
Single equation testing shows lack of evidences in favor of the theoretical hypothesis. Therefore, panel unit root or cointegration testing has been proposed as a solution to the empirical puzzle. Indeed, the higher power of the panel tests has led to the rejection of the null of absence of PPP hypothesis. Nevertheless, many studies have applied panel tests on groups of countries - usually not related with each other - but the conclusion they draw could be only applied to the group as a whole, without giving directions on which countries follow the theoretical prediction and which do not. Therefore, it seems appropriate to analyze this with the *SUR-bounds* test which can provide individual tests treating cross-country dependency at the same time.
The theory states that nominal exchange rate and price ratio of two countries should move together in the long-run. Hence, the long-run cointegrating relation is:

$$s_{it} = \alpha_i + \beta_i p_{it} + e_{it}$$

where s_{it} is the log nominal exchange rate of country i with respect to the US and p_{it} is the log price differencial between country i and the US. Here, the US are taken as benchmark country. The data source is the OECD database. Differently from the strong PPP hypothesis which postulate slope parameters β_i equal to one, the weak version of the theory only requires that the two series are cointegrated.
The individual ECM testing regressions are:

$$\Delta s_{it} = \pi_{is} s_{i,t-1} + \pi_{ip} p_{i,t-1} + \sum_{j=1}^{p_i-1} \psi_{isj} \Delta s_{i,t-j} + \sum_{j=0}^{q_i-1} \psi_{ipj} \Delta p_{i,t-j} + u_{it}$$

i=aus,bel,de,den,fin,fra,gr,ire,ita,lux,nl,por,spa,swe,uk

Under the null hypothesis of no cointegration, i.e. $H_0^i : \pi_{is} = 0 \ \pi_{ip} = 0$, the PPP hypothesis does not hold.
For the sample under analysis, the correlations between the error series u_{it} are remarkable. As the following correlation matrix shows, they range between 0.60 and 0.99. Therefore, the power gains are expected to be important.

$$\hat{\Sigma} = \begin{array}{r} \text{aut} \\ \text{bel} \\ \text{de} \\ \text{den} \\ \text{fin} \\ \text{fra} \\ \text{gr} \\ \text{ire} \\ \text{ita} \\ \text{lux} \\ \text{nl} \\ \text{por} \\ \text{spa} \\ \text{swe} \\ \text{uk} \end{array} \begin{bmatrix} 1.00 \\ 0.96 & 1.00 \\ 0.99 & 0.97 & 1.00 \\ 0.96 & 0.95 & 0.96 & 1.00 \\ 0.76 & 0.73 & 0.75 & 0.78 & 1.00 \\ 0.91 & 0.87 & 0.91 & 0.90 & 0.78 & 1.00 \\ 0.72 & 0.70 & 0.73 & 0.71 & 0.62 & 0.71 & 1.00 \\ 0.85 & 0.82 & 0.84 & 0.84 & 0.77 & 0.82 & 0.69 & 1.00 \\ 0.82 & 0.80 & 0.82 & 0.80 & 0.77 & 0.83 & 0.67 & 0.80 & 1.00 \\ 0.97 & 0.99 & 0.97 & 0.96 & 0.73 & 0.87 & 0.68 & 0.82 & 0.80 & 1.00 \\ 0.98 & 0.97 & 0.99 & 0.97 & 0.77 & 0.90 & 0.73 & 0.84 & 0.81 & 0.97 & 1.00 \\ 0.86 & 0.84 & 0.86 & 0.86 & 0.75 & 0.81 & 0.69 & 0.76 & 0.79 & 0.84 & 0.85 & 1.00 \\ 0.71 & 0.71 & 0.71 & 0.74 & 0.88 & 0.74 & 0.58 & 0.72 & 0.78 & 0.71 & 0.72 & 0.71 & 1.00 \\ 0.83 & 0.80 & 0.82 & 0.81 & 0.76 & 0.84 & 0.67 & 0.77 & 0.74 & 0.80 & 0.83 & 0.74 & 0.69 & 1.00 \\ 0.62 & 0.60 & 0.62 & 0.63 & 0.71 & 0.62 & 0.69 & 0.74 & 0.67 & 0.59 & 0.64 & 0.63 & 0.65 & 0.67 & 1.00 \end{bmatrix}$$

Table 9 presents the results for the PSS test and for the *SUR-bounds* test applied to the individual countries and to the panel as a whole.
While the PSS test is only rejecting the null of no cointegration - i.e. no PPP hypothesis - for Sweden and France, the *SUR-bounds* individual tests provide evidence in favor of the PPP hypothesis for almost all the countries, with the only exception of UK and Greece, as well as the panel statistics.
Thus, differently for the PSS single equation test, the *SUR-bounds* test empirically supports the weak-form PPP hypothesis and this is basically due to the increased power of the new methodology with respect to the PSS test, as shown in the previous sections.

6 Conclusions

This paper proposes a new test for panel cointegration. It extends the bounds test of Pesaran et al. (2001) to a group of cross-sectional units by considering the testing regressions in a *SUR* system in order to take into account the cross correlation in the error terms and improve the power properties.
It tests unit-specific null hypotheses of no significance of the autoregressive parameters in the conditional ECM models. If the null is rejected, there is no level relationship among the variables.
Differently from the other panel cointegration tests, the test proposed herein provides unit specific tests as well as panel tests. This allows to identify which units of the panel are cointegrated and which are not.

The algorithm to implement the test is illustrated and a Monte Carlo experiment is performed in order to check the size and power properties. Many scenarios are simulated, according to the degree of cross-section correlation,

the number of units in the panel and the number of observations.
The new test shows minimal or absent size distortion and substantial gains on the power side over the Pesaran et al. (2001) test.
The comparison with Westerlund (2007) test favours the *SUR-bounds* test when the cross sectional correlation is medium or high.
The empirical application on PPP hypothesis testing, illustrates how the increased power properties allow to reject the null of absence of cointegration between the exchange rate and the price differential for a panel of the EU15 countries and it confirms the validity of the PPP hypothesis.

Table 1: *PSS* test 0.05 critical values

T	**F**		**W**		**t**	
50	3.325	4.284	6.650	8.568	-3.054	-2.172
100	3.213	4.177	6.427	8.354	-2.518	-2.302
200	3.134	4.044	6.268	8.089	-2.446	-2.261

Notes: The critical values of joint and of the t statistics are respectively the 95th and the 5th percentile of the empirical distribution obtained with 10,000 replications.

Table 2: *SUR-bounds* test 0.05 critical values

Correlation	***N***	***T***	F_i		W_i		t_i		***sys-F***		***sys-W***	
L	5	50	3.858	4.926	7.972	10.199	-2.176	-2.915	2.361	2.984	24.259	31.188
		100	3.507	4.596	7.132	9.348	-2.096	-2.783	2.141	2.755	21.579	28.134
		200	3.330	4.358	6.717	8.792	-2.047	-2.699	2.042	2.632	20.495	26.748
	10	50	4.583	5.767	9.423	11.887	-2.358	-3.162	2.467	3.053	50.133	62.747
		100	3.851	5.008	7.817	10.171	-2.212	-2.914	2.016	2.581	40.870	52.331
		200	3.561	4.636	7.175	9.343	-2.145	-2.792	1.828	2.379	36.733	47.916
	20	50	6.414	7.822	13.089	16.031	-2.744	-3.716	3.188	3.836	129.515	157.719
		100	4.651	5.866	9.416	11.892	-2.434	-3.168	2.142	2.675	86.722	108.457
		200	3.950	5.143	7.952	10.358	-2.269	-2.946	1.803	2.280	72.535	91.899
M	5	50	3.988	5.209	8.110	10.657	-2.256	-2.985	2.375	2.955	24.115	30.298
		100	3.666	4.889	7.391	9.882	-2.198	-2.875	2.129	2.702	21.416	27.344
		200	3.509	4.672	7.048	9.391	-2.174	-2.811	2.046	2.610	20.536	26.214
	10	50	4.826	6.264	9.762	12.758	-2.488	-3.282	2.491	3.008	50.424	61.415
		100	4.139	5.538	8.327	11.154	-2.380	-3.060	2.056	2.556	41.326	51.515
		200	3.860	5.197	7.737	10.430	-2.331	-2.953	1.854	2.342	37.165	47.019
	20	50	7.055	8.875	14.224	17.991	-2.923	-3.945	3.368	3.944	135.419	160.354
		100	5.072	6.736	10.173	13.534	-2.644	-3.370	2.171	2.655	87.137	106.877
		200	4.421	6.043	8.852	12.108	-2.529	-3.163	1.832	2.269	73.333	90.986
H	5	50	4.186	5.696	8.191	11.287	-2.359	-3.091	2.355	2.847	23.243	28.348
		100	3.896	5.430	7.711	10.787	-2.347	-3.023	2.165	2.647	21.176	26.094
		200	3.786	5.300	7.520	10.550	-2.348	-3.000	2.056	2.518	20.571	25.183
	10	50	5.425	7.396	10.563	14.554	-2.721	-3.526	2.580	3.009	49.897	58.791
		100	4.804	6.702	9.432	13.204	-2.679	-3.338	2.068	2.492	40.963	49.556
		200	4.474	6.389	8.851	12.653	-2.634	-3.254	1.862	2.288	36.764	45.270
	20	50	9.004	11.639	17.376	22.818	-3.324	-4.468	3.880	4.235	149.188	165.323
		100	6.360	8.800	12.347	17.168	-3.092	-3.810	2.289	2.705	88.826	104.269
		200	5.814	8.331	11.432	16.345	-3.081	-3657	1.854	2.225	72.811	87.46

Notes: The critical values of joint and of the t statistics are respectively the 95th and the 5th percentile of the empirical distribution obtained with 10,000 replications. "Correlation" refers to the degree (L=low,M=medium,H=high) of residuals' correlation of the panel. Average critical values are reported.

Table 3: Size properties at nominal size 0.05

Correlation	***N***	***T***	F_i		W_i		t_i		***sys-F***	***sys-W***
			SUR	***PSS***	***SUR***	***PSS***	***SUR***	***PSS***	***SUR***	***SUR***
L	5	50	0.051	0.055	0.060	0.055	0.060	0.058	0.043	0.077
		100	0.047	0.053	0.052	0.053	0.054	0.050	0.056	0.057
		200	0.051	0.057	0.054	0.057	0.055	0.054	0.050	0.056
	10	50	0.049	0.054	0.058	0.054	0.055	0.055	0.041	0.068
		100	0.050	0.050	0.054	0.050	0.053	0.051	0.052	0.071
		200	0.052	0.052	0.054	0.052	0.053	0.051	0.059	0.063
	20	50	0.049	0.051	0.054	0.051	0.051	0.051	0.036	0.045
		100	0.051	0.051	0.054	0.051	0.053	0.046	0.056	0.070
		200	0.051	0.046	0.053	0.046	0.054	0.047	0.056	0.067
M	5	50	0.048	0.055	0.054	0.055	0.051	0.059	0.048	0.064
		100	0.048	0.051	0.050	0.051	0.049	0.049	0.048	0.052
		200	0.050	0.056	0.051	0.056	0.047	0.052	0.052	0.053
	10	50	0.048	0.055	0.052	0.055	0.050	0.054	0.049	0.055
		100	0.047	0.050	0.049	0.050	0.051	0.052	0.053	0.058
		200	0.051	0.052	0.051	0.052	0.051	0.051	0.055	0.057
	20	50	0.045	0.050	0.047	0.050	0.046	0.050	0.038	0.039
		100	0.049	0.054	0.050	0.054	0.050	0.051	0.054	0.057
		200	0.051	0.045	0.051	0.045	0.052	0.048	0.049	0.051
H	5	50	0.048	0.058	0.047	0.058	0.044	0.060	0.049	0.043
		100	0.044	0.053	0.041	0.053	0.044	0.052	0.049	0.046
		200	0.046	0.054	0.044	0.054	0.045	0.051	0.050	0.046
	10	50	0.044	0.056	0.038	0.056	0.038	0.056	0.038	0.030
		100	0.044	0.051	0.039	0.051	0.041	0.053	0.040	0.027
		200	0.047	0.053	0.044	0.053	0.046	0.053	0.048	0.044
	20	50	0.034	0.050	0.028	0.050	0.029	0.052	0.019	0.011
		100	0.045	0.050	0.038	0.050	0.039	0.047	0.043	0.027
		200	0.047	0.045	0.042	0.045	0.043	0.050	0.047	0.032

Notes: For the individual statistics F_i, W_i and t_i average size (over N) of PSS and SUR test are reported. Size is based on 2000 replications.

Table 4: Power properties at nominal size 0.05 - α_i=U(-0.01,-0.03)

Correlation	***N***	***T***	F_i		W_i		t_i		***sys-F***	***sys-W***
			SUR	***PSS***	***SUR***	***PSS***	***SUR***	***PSS***	***SUR***	***SUR***
L	5	50	0.088	0.084	0.103	0.084	0.092	0.084	0.142	0.174
		100	0.133	0.125	0.142	0.125	0.118	0.104	0.302	0.329
		200	0.280	0.251	0.286	0.251	0.254	0.220	0.690	0.699
	10	50	0.084	0.082	0.097	0.083	0.086	0.078	0.153	0.197
		100	0.144	0.121	0.152	0.121	0.125	0.107	0.465	0.503
		200	0.281	0.243	0.288	0.243	0.249	0.216	0.906	0.913
	20	50	0.080	0.076	0.089	0.076	0.078	0.073	0.131	0.161
		100	0.142	0.124	0.149	0.124	0.124	0.101	0.643	0.685
		200	0.277	0.223	0.282	0.223	0.251	0.210	0.995	0.995
M	5	50	0.096	0.080	0.107	0.080	0.093	0.086	0.175	0.201
		100	0.177	0.125	0.183	0.125	0.144	0.105	0.397	0.409
		200	0.364	0.254	0.367	0.254	0.313	0.220	0.785	0.789
	10	50	0.103	0.086	0.109	0.860	0.091	0.080	0.221	0.242
		100	0.190	0.122	0.194	0.122	0.148	0.106	0.606	0.619
		200	0.379	0.246	0.382	0.246	0.323	0.217	0.964	0.964
	20	50	0.090	0.075	0.093	0.075	0.075	0.072	0.176	0.188
		100	0.186	0.130	0.188	0.130	0.142	0.113	0.832	0.836
		200	0.374	0.228	0.375	0.228	0.326	0.209	1.000	1.000
H	5	50	0.121	0.079	0.114	0.079	0.087	0.083	0.270	0.254
		100	0.252	0.124	0.244	0.124	0.178	0.108	0.596	0.585
		200	0.514	0.251	0.506	0.251	0.426	0.216	0.927	0.921
	10	50	0.129	0.085	0.116	0.085	0.081	0.079	0.373	0.336
		100	0.272	0.126	0.256	0.126	0.179	0.110	0.849	0.812
		200	0.541	0.248	0.530	0.248	0.449	0.218	0.997	0.996
	20	50	0.097	0.074	0.084	0.074	0.055	0.071	0.350	0.292
		100	0.265	0.122	0.244	0.122	0.164	0.102	0.989	0.979
		200	0.519	0.229	0.503	0.229	0.435	0.209	1.000	1.000

Notes: For the individual statistics F_i, W_i and t_i average power (over N) of PSS and SUR test are reported. Power is based on 2000 replications.

Table 5: Power properties at nominal size 0.05 - α_i=U(-0.07,-0.09)

Correlation	***N***	***T***	F_i		W_i		t_i		***sys-F***	***sys-W***
			SUR	***PSS***	***SUR***	***PSS***	***SUR***	***PSS***	***SUR***	***SUR***
L	5	50	0.232	0.216	0.262	0.216	0.260	0.234	0.592	0.650
		100	0.507	0.468	0.525	0.468	0.559	0.508	0.968	0.974
		200	0.886	0.857	0.891	0.857	0.926	0.903	1.000	1.000
	10	50	0.246	0.223	0.271	0.223	0.261	0.230	0.785	0.838
		100	0.532	0.467	0.547	0.467	0.565	0.514	1.000	1.000
		200	0.888	0.854	0.893	0.854	0.925	0.901	1.000	1.000
	20	50	0.241	0.211	0.257	0.211	0.240	0.219	0.861	0.900
		100	0.535	0.470	0.547	0.470	0.560	0.496	1.000	1.000
		200	0.890	0.834	0.893	0.834	0.921	0.895	1.000	1.000
M	5	50	0.301	0.215	0.322	0.215	0.309	0.231	0.690	0.713
		100	0.638	0.462	0.646	0.462	0.662	0.504	0.985	0.986
		200	0.951	0.855	0.954	0.855	0.965	0.902	1.000	1.000
	10	50	0.327	0.224	0.339	0.224	0.317	0.233	0.889	0.903
		100	0.667	0.470	0.673	0.470	0.684	0.515	0.995	1.000
		200	0.955	0.852	0.956	0.852	0.967	0.901	1.000	1.000
	20	50	0.305	0.210	0.312	0.210	0.281	0.218	0.946	0.948
		100	0.665	0.483	0.669	0.483	0.678	0.529	1.000	1.000
		200	0.955	0.833	0.956	0.833	0.968	0.894	1.000	1.000
H	5	50	0.455	0.213	0.444	0.213	0.401	0.229	0.838	0.823
		100	0.828	0.465	0.822	0.465	0.820	0.505	0.997	0.997
		200	0.993	0.851	0.992	0.851	0.993	0.898	1.000	1.000
	10	50	0.480	0.227	0.456	0.227	0.410	0.234	0.975	0.965
		100	0.857	0.471	0.846	0.471	0.844	0.515	1.000	1.000
		200	0.995	0.853	0.994	0.853	0.996	0.903	1.000	1.000
	20	50	0.424	0.207	0.393	0.207	0.339	0.218	0.993	0.988
		100	0.865	0.476	0.850	0.476	0.851	0.501	1.000	1.000
		200	0.996	0.834	0.995	0.834	0.997	0.894	1.000	1.000

Notes: For the individual statistics F_i, W_i and t_i average power (over N) of PSS and SUR test are reported. Power is based on 2000 replications.

Table 6: Serial correlation in the error term

	CS	F_i		W_i		t_i		***sys-F***	***sys-W***
		SUR	***PSS***	***SUR***	***PSS***	***SUR***	***PSS***	***SUR***	***SUR***
	L	0.050	0.050	0.054	0.050	0.053	0.051	0.052	0.071
no corr	M	0.047	0.050	0.049	0.050	0.051	0.052	0.053	0.058
	H	0.044	0.051	0.039	0.051	0.041	0.053	0.040	0.027
	L	0.051	0.055	0.060	0.055	0.060	0.058	0.043	0.077
$MA(1)$	M	0.050	0.050	0.051	0.050	0.047	0.050	0.058	0.062
	H	0.053	0.054	0.058	0.054	0.055	0.053	0.058	0.072
	L	0.075	0.081	0.084	0.081	0.053	0.055	0.132	0.183
$MA(2)$	M	0.069	0.080	0.073	0.080	0.047	0.057	0.130	0.144
	H	0.042	0.054	0.038	0.054	0.037	0.052	0.053	0.039

Notes: For the individual statistics F_i, W_i and t_i average size (over N) of PSS and SUR test are reported. Size is based on 2000 replications and N and T are set to 10 and 100, respectively.

Table 7: Mixed panel

Correlation	***N***	F_i		W_i		t_i		***sys-F***	***sys-W***
		H_0	H_1	H_0	H_1	H_0	H_1		
L	1	0.028	0.442	0.028	0.448	0.030	0.465	0.984	0.986
M	1	0.019	0.596	0.019	0.595	0.019	0.602	0.998	0.998
H	1	0.009	0.822	0.009	0.809	0.006	0.803	1.000	1.000
L	3	0.030	0.441	0.030	0.446	0.032	0.463	0.924	0.931
M	3	0.022	0.598	0.021	0.597	0.022	0.603	0.984	0.983
H	3	0.012	0.823	0.011	0.811	0.009	0.802	1.000	1.000
L	5	0.029	0.443	0.030	0.449	0.032	0.463	0.762	0.777
M	5	0.023	0.601	0.023	0.601	0.024	0.608	0.922	0.919
H	5	0.012	0.826	0.011	0.814	0.012	0.801	0.995	0.993
L	7	0.030	0.439	0.031	0.445	0.033	0.464	0.416	0.435
M	7	0.025	0.607	0.025	0.606	0.026	0.609	0.622	0.613
H	7	0.015	0.825	0.013	0.813	0.014	0.797	0.922	0.913
L	9	0.030	0.432	0.032	0.437	0.034	0.452	0.146	0.144
M	9	0.029	0.609	0.029	0.609	0.030	0.612	0.027	0.025
H	9	0.023	0.835	0.020	0.823	0.022	0.804	0.319	0.299

Notes: Columns H_0 report average power for the n not cointegrated units while columns H_1 report average power for the $N-n$ cointegrated units. Power is based on 2000 replications and N and T are set to 10 and 100, respectively.

Table 8: Comparison with Westerlund (2007)

		SUR-bounds test					***Westerlund (2007)***			
	CS	F_i	W_i	t_i	***sys-F***	***sys-W***	G_τ	G_α	P_τ	P_α
	L	0.049	0.062	0.058	0.052	0.089	0.072	0.019	0.062	0.028
size	M	0.043	0.049	0.047	0.040	0.053	0.076	0.021	0.069	0.031
	H	0.033	0.027	0.027	0.034	0.025	0.073	0.032	0.076	0.041
	L	0.143	0.119	0.121	0.364	0.477	0.486	0.146	0.557	0.371
power -0.03,-0.01	M	0.161	0.174	0.133	0.485	0.536	0.427	0.140	0.507	0.318
	H	0.226	0.202	0.130	0.753	0.677	0.333	0.123	0.416	0.235
	L	0.411	0.455	0.478	0.992	0.997	1.000	0.991	1.000	1.000
power -0.09,-0.07	M	0.547	0.571	0.576	0.998	0.999	0.993	0.949	1.000	0.999
	H	0.757	0.730	0.717	1.000	1.000	0.962	0.848	0.992	0.986

Notes: Size and power are based on 2000 replications and N and T are set to 10 and 100, respectively.

Table 9: Test statistics for the verification of the PPP hypothesis

Country	$\mathbf{W_i - PSS}$		$\mathbf{W_i - SUR}$		***cv*** $\mathbf{H_0}$	***cv*** $\mathbf{H_1}$	$\mathbf{t_i - PSS}$		$\mathbf{t_i - SUR}$		***cv*** $\mathbf{H_0}$	***cv*** $\mathbf{H_1}$
aut	3.24		58.30	*	15.16	17.24	-1.48		-7.20	*	-3.92	-4.30
bel	8.45		49.88	*	11.36	13.93	-2.06		-6.04	*	-3.72	-4.07
de	4.03		60.88	*	11.60	13.62	-1.83		-7.28	*	-2.93	-3.25
den	3.58		48.33	*	13.30	15.21	-0.86		-6.42	*	-4.00	-4.27
fin	4.09		36.47	*	10.07	11.64	-2.01		-6.03	*	-3.90	-4.20
fra	14.01	*	60.36	*	13.74	17.04	-3.14		-7.34	*	-3.61	-3.98
gr	2.94		8.66		15.55	16.99	-1.51		-2.90		-3.13	-3.51
ire	7.84		30.07	*	14.92	17.47	-2.80		-4.85	*	-3.59	-3.97
ita	6.97		34.73	*	12.11	14.35	-2.50		-5.72	*	-3.66	-3.90
lux	7.94		49.28	*	15.00	17.84	-2.04		-5.98	*	-3.48	-3.92
nl	3.81		59.21	*	17.32	20.16	-1.60		-6.92	*	-2.78	-3.07
por	4.15		23.97	*	18.00	20.29	-1.37		-3.69	*	-3.36	-3.68
spa	4.21		21.42	*	10.76	12.89	-1.91		-4.52	*	-3.00	-3.40
swe	13.56	*	35.70	*	15.73	18.41	-3.78	#	-5.86	*	-3.10	-3.51
UK	6.14		13.11		17.01	19.96	-2.27		-3.61	#	-3.58	-3.93
			$sys - W$		***cv*** $\mathbf{H_0}$	***cv*** $\mathbf{H_1}$						
	Joint test		114.67		80.59	94.02						

Notes: * indicates rejection of the null hypothesis of no-cointegration (i.e. no-PPP hypothesis). # indicates the cases in which the statistic falls between the two bounds.The sample is constituted by quarterly data, from 1976:1 to 1998:3. Data source is OECD database.

References

Banerjee, A., Dolado, J., and Mestre, R. (1998). Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of Time Series Analysis*, 19(3):267–283.

Banerjee, A., Marcellino, M., and Osbat, C. (2004). Some cautions on the use of panel methods for integrated series of macroeconomic data. *The Econometrics Journal*, 7(2):322–340.

Breitung, J. (2005). A parametric approach to the estimation of cointegration vectors in panel data. *Econometric Reviews*, 24(2):151–173.

Breitung, J. and Das, S. (2005). Panel unit root tests under cross-sectional dependence. *Statistica Neerlandica*, 59(4):414–433.

Breitung, J. and Pesaran, M. (2008). Unit roots and cointegration in panels. *The Econometrics of Panel Data*, pages 279–322.

Breuer, J., McNown, R., and Wallace, M. (2002). Series-specific unit root tests with panel data. *Oxford Bulletin of Economics and Statistics*, 64(5):527–546.

Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics*, 120(2):263 – 293.

Chong, Y., Jordà, Ò., and Taylor, A. M. (2012). The harrod–balassa–samuelson hypothesis: real exchange rates and their long-run equilibrium. *International Economic Review*, 53(2):609–634.

Gengenbach, C., Palm, F. C., and Urbain, J.-P. (2006). Cointegration testing in panels with common factors. *Oxford Bulletin of Economics and Statistics*, 68(s1):683–719.

Gengenbach, C., Urbain, J.-P., and Westerlund, J. (2008). Panel error correction testing with global stochastic trends. Research Memoranda 051, Maastricht : METEOR, Maastricht Research School of Economics of Technology and Organization.

Herzer, D., Strulik, H., and Vollmer, S. (2012). The long-run determinants of fertility: one century of demographic change 1900-1999. *Journal of Economic Growth*, 17(4):357–385.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–80.

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, 90(1):1–44.

Kónya, L. (2006). Exports and growth: Granger causality analysis on oecd countries with a panel data approach. *Economic Modelling*, 23(6):978 – 992.

Larsson, R., Lyhagen, J., and LÃűthgren, M. (2001). Likelihood-based cointegration tests in heterogeneous panels. *Econometrics Journal*, 4(1):109–142.

Mark, N. C., Ogaki, M., and Sul, D. (2005). Dynamic seemingly unrelated cointegrating regressions. *Review of Economic Studies*, 72(3):797–820.

Morley, B. (2006). Causality between economic growth and immigration: An ardl bounds testing approach. *Economics Letters*, 90(1):72 – 76.

Narayan, P. K. (2005). The saving and investment nexus for china: evidence from cointegration tests. *Applied Economics*, 37(17):1979–1990.

Narayan, P. K. and Smyth, R. (2004). Crime rates, male youth unemployment and real income in australia: evidence from granger causality tests. *Applied Economics*, 36(18):2079–2095.

O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics*, 44(1):1–19.

Pedroni, P. (2004). Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the ppp hypothesis. *Econometric theory*, 20(03):597–625.

Pesaran, M. and Shin, Y. (1998). An autoregressive distributed-lag modelling approach to cointegration analysis. *Econometric Society Monographs*, 31:371–413.

Pesaran, M. and Shin, Y. (2002). Long-run structural modelling. *Econometric Reviews*, 21(1):49–87.

Pesaran, M., Shin, Y., and Smith, R. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16(3):289–326.

Pesavento, E. (2004). Analytical evaluation of the power of tests for the absence of cointegration. *Journal of Econometrics*, 122(2):349–384.

Rassenfosse, G. d. and Potterie, B. v. P. d. l. (2012). On the price elasticity of demand for patents*. *Oxford Bulletin of Economics and Statistics*, 74(1):58–77.

Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. *Journal of Business & Economic Statistics*, 7(2):147–59.

Westerlund, J. (2007). Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, 69(6):709–748.

Westerlund, J. (2008). Panel cointegration tests of the fisher effect. *Journal of Applied Econometrics*, 23(2):193–233.

Westerlund, J. and Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks. *Oxford Bulletin of Economics and Statistics*, 70(5):665–704.

Zivot, E. (2000). The power of single equation tests for cointegration when the cointegrating vector is prespecified. *Econometric Theory*, 16(03):407–439.

The Expectation Hypothesis of the Term Structure of Very Short-Term Rates: Evidence from a New Testing Approach

Vanessa Gunnella[1]

Abstract

This paper empirically tests the Expectation Hypothesis of the term structure of the US repurchasing agreements (repo) rates, considered in a Vector Auto Regression (VAR) model. A multiple hypotheses approach is adopted, in order to jointly test all the statistical hypotheses implied by the EH, i.e. the long-run and short-run implications of the theory. Furthermore, the testing procedures are carried out by taking into account heteroskedasticity through bootstrap inference, White correction and rolling windows analysis. Differently from previous results, overall evidence in favor of the statistical non-rejection of the EH is found. In particular, the rolling window analysis clarifies that the EH has been rejected only during periods of turmoil of the financial/repo markets.

JEL C12 C14 C18 C32 E43 G14
Keywords Expectation Hypothesis, Repo, VAR Testing, Heteroskedasticity, Multiple Hypothesis Testing

[1]University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Email: vanessa.gunnella2@unibo.it
I wish to thank James Hamilton, Fabrizio Iacone, Anders Rahbek, Michel Van Der Wel and all the participants of the PhD Forum - Department of Economics - University of Bologna, Economics Thursday Workshop - University of York, the BMRC-DEMS 2014 Conference, the IAAE 2014 Conference, the CFE 2014 and the ICEEE2015 Conference where this version and previous versions of the paper have been presented. A special thank goes to my supervisor Luca Fanelli, for his constant guidance and encouragement. I also gratefully acknowledge the support of IAAE in the form of a grant award for the presentation of this paper at the IAAE 2014 Conference.

1 Introduction

The aim of this paper is to test the Expectation Hypothesis of the Term Structure (EHTS) of the US repo rates. It contributes to the literature by taking into account cointegration and changes in volatility over the sample. The testing procedure is carried out in a multiple hypothesis framework, which means that the long-run and the short-run implications of the theory are addressed jointly through a strict size control.
The repo (repurchase agreement) contracts are among the most relevant financial instruments in terms of trading volumes. It is a form of collateralized debt in which the collateral is often represented by Treasury or Government Bonds, because of their high liquidity. The repo market constitutes the main channel of short-term liquidity provision for financial institutions and serves the cash storage needs of hedge founds (see Adrian and Shin, 2010). Its market size in the US amounts to about $10 trillion, according to estimates by the Task Force on Tri-Party Repo Infrastructure and Bank of International Settlement (see Gorton and Metrick, 2012). Moreover, repo rates reflect the liquidity and collateral value of the assets, and the repo-Libor spread is considered by financial agents a good proxy for liquidity.
Hence, the repo market and its functioning play a crucial role in the money market. It is thus relevant to assess whether the repo market is well operating, namely, whether long-term interest rates correctly reflect future expectations on short-term interest rate, as the EHTS predicts.

From the empirical point of view, the verification of the EHTS constitutes a controversial topic. While EHTS represents a commonly accepted framework in economic theory, the econometric evidence is mixed. The implications of the EHTS are generally rejected when tested on US bonds term structure data (see Campbell and Shiller, 1987 and Bekaert and Hodrick, 2001 among the others).
More specifically, as regards the EHTS of the US repo Term Structure, Longstaff (2000) finds empirical support to the EHTS for very short-term repo rates, whereas Della Corte et al. (2008) statistically reject the hypothesis in an extension of Longstaff (2000) data set. Della Corte et al. (2008) apply Campbell and Shiller (1987) methodology as revised by Bekaert and Hodrick (2001).
Existing approaches suffer from some drawbacks. Even though Della Corte et al. (2008) improve upon Longstaff (2000)'s single equation framework by increasing the power of the testing procedure with a VAR approach, they do not exploit the non-stationarity of repo rates - hence cointegration analysis - and they treat heteroskedasticity parametrically, that is by simulating a sample with $GARCH$ innovations with the bootstrap.

Therefore, the aim of this paper is to improve on the econometric and empirical analysis along the following dimension:

1. An "inexact" formulation of the EHTS is considered. In the inexact form, the stochastic disturbance which affects the relationships between interest rates at different maturities is consistent with a time-varying component in the risk premia. The time-varying component represents transitory deviations from the equilibrium.

2. The time series properties of the interest rates are seriously taken into account by

treating them as a non-stationary cointegrated system. While Della Corte et al. (2008) correct for small sample bias and apply Campbell and Shiller's (1987) method for stationary data, this paper builds on Campbell and Shiller (1987) "two steps" procedure which is devised for non-stationary time series. The approach is further extended by modeling all the interest rates of the term structure as generated by a joint process in order to make a more realistic assumption about the Data Generating Process[2].

3. The effect of the heteroskedasticity is controlled by means of testing procedures which are robust to variance and covariance shifts over time.
   On the one hand, Cavaliere et al. (2012, 2014) and Boswijk et al. (2013) methodology for heteroskedastic co-integrated VAR models and Hafner and Herwartz (2009) methodology for heteroskedastic stationary VAR model are applied. On the other, rolling window analysis is performed in order to assume a more general type of parameters variation which not only involves the conditional and unconditional variance, but also the other parameters.
   It is important to notice that, by introducing heteroskedasticity control, this paper contributes to the general literature of present value models testing[3], because many (if not most) applications are characterized by time-varying volatility and therefore should use the approach proposed herein. For instance, it is well established that financial time series are affected by heteroskedasticity.

4. The implications of the EHTS at different frequencies are tested jointly and not separately. Therefore, the size of the testing procedure is strictly controlled.
   Testing for the expectation hypothesis requires that both the long-run and the short-run properties are verified. For this reason, it is important to test both properties jointly at a pre-fixed level of significance. The type I error should be such that it does not exceed the sum of the type I errors pre-fixed for each test involved.

5. Through the rolling window analysis, the paper introduces the idea that the assessment of EHTS is time-dependent. Indeed, it is reasonable to expect that the assumptions behind the EHTS are not always holding throughout the sample.

The statistical testing procedures lead to the non-rejection of the long-run implications of EHTS. For what regards short-term implications, the restrictions are not rejected in the inexact form of EHTS. In particular, this result is due to the size correction. Therefore, considered as a whole, the EHTS is not rejected.
Furthermore, the rolling window analysis clarifies that the EHTS is only rejected in periods of turbulence of financial markets.

The remainder of the paper is organized as follow: Section 2 introduces the theoretical model. Section 3 describes the econometric methods applied on the data

[2]Sarno et al. (2007) applied the Bekaert-Hodrick methodology to all possible trivariate combinations of their 12 bond yields.

[3]Recent empirical contributions which use Campbell and Shiller (1987) methodology to test present value models are: Sbordone (2005) and Fanelli (2008) for what regards New Keynesian Phillips Curve; Engel and West (2005) in an application on exchange rates and Campa and Gavilan (2011) for the empirical verification of the Permanent Income Hypothesis.

for the term structure of Repo rates (illustrated in Section 4). The results of the empirical analysis are displayed and discussed in Section 5. Section 6 offers some conclusions.

2 Theoretical framework

The EHTS of interest rates states that the return on the h-period investment should equal the return of a 1 period investment rolled over h times, plus a term premium:

$$(1 + R_t^{(h)})^h = \Theta_t^{(h)} \prod_{j=0}^{h-1} (1 + E_t r_{t+j})$$

where $R_t^{(h)}$ is the annual interest rate on a h period investment, r_t is the annual return on a 1 period investment, $E_t = E(\, . \mid I_t)$ is the expected value conditional on the information set I_t and $\Theta_t^{(h)}$ is the term premium.
Taking logs, the relation becomes:

$$R_t^{(h)} = \frac{1}{h} \sum_{j=0}^{h-1} E_t r_{t+j} + \theta_t^{(h)} \qquad (1)$$

where $\theta_t^{(h)} = (1/h) log(\Theta_t^{(h)})$.
The term premium is modeled such that it is the sum of a constant long-run component, $\bar{\theta}^{(h)}$ and a time varying component, $\tilde{\theta}_t^{(h)}$:[4]

$$\theta_t^{(h)} = \bar{\theta}^{(h)} + \tilde{\theta}_t^{(h)}$$

The time varying component $\tilde{\theta}_t^{(h)}$ can be interpreted as a term which captures what the EHTS is not able to explain, *i.e.* observable transitory deviations from the equilibrium conditions (*i.e.* variations of expected equilibrium returns, change in the risk free rate and riskiness, etc.). In principle, it can be assumed to follow several stochastic processes. However, if $\tilde{\theta}_t^{(h)}$ captures transitory deviations from the equilibrium conditions, its time series pattern cannot exhibit too much persistence and this term must be at least stationary. In the empirical analysis, it will be assumed that $\tilde{\theta}_t^{(h)}$ obeys a Martingale Difference Sequence (MDS), *i.e.* a process such that $E[\tilde{\theta}_t^{(h)} \mid I_{t-1}] = 0$. This means that the process is not forecastable with currently available information and hence there is no room for arbitrage.
As González and Gonzalo (2000) show, adding $\tilde{\theta}_t^{(h)}$ does not change the equilibrium conditions from which the model is derived, in the sense that the model with time-varying term-premium is derived from the same non-arbitrage conditions of the exact present value model.

[4] Campbell and Shiller (1987) and Hansen (2003) among others, consider a constant term premium in the theoretical model, while González and Gonzalo (2000) by assuming a time-varying term premium - but abstracting from the constant part - derive testing procedures for the "inexact" present value model.

Subtracting r_t from both sides of equation (1), an equation in terms of the *spread* $S_t^{(h)} := R_t^{(h)} - r_t$ is obtained:

$$S_t^{(h)} = \sum_{j=0}^{h-1} \Big(1 - \frac{j}{h}\Big) E_t \Delta r_{t+j} + \theta_t^{(h)}$$

Shiller (1979) derived an approximation of the equation above from a linearization of an expectations model and obtained the following solution through recursive substitution:

$$S_t^{(h)} = \sum_{j=0}^{h-1} \delta_h^j E_t \Delta r_{t+j} + \theta_t^{(h)}$$

with $\delta_h := 1/(1+R)$ being a parameter of linearization.
By writing out the term premium, the following equation is obtained:

$$S_t^{(h)} = \sum_{j=0}^{h-1} \delta_h^j E_t \Delta r_{t+j} + \bar{\theta}^{(h)} + \tilde{\theta}_t^{(h)} \tag{2}$$

An equivalent parametrization of equation (2) is derived after some algebraic manipulations which lead to the expression:

$$S_t^{(h)} = \delta_h E_t S_{t+1}^{(h)} + \delta_h E_t \Delta r_{t+1} + \theta_t^{(h)} \tag{3}$$

i.e. a *inexact* formulation of the Present Value model (see González and Gonzalo, 2000 and Fanelli, 2008).

3 Methodology

Provided that the interest rates are integrated of order one ($I(1)$) variables, the EHTS can be tested following Campbell and Shiller's (1987) approach.
The idea is to nest the model in equation (2) within a VAR model for the interest rates. The testing procedure consists in two steps.

The authors outline a testable implication of the theoretical equations above, which is verified in the first step. If the short-term rate r_t is an $I(1)$ variable, the term $\sum_{j=0}^{h-1} \delta_h^j E_t \Delta r_{t+j}$ in equation (2) is stationary, as well as $\tilde{\theta}_t^{(h)}$, if it is assumed to be a MDS or an $MA(q)$ process, for instance. It follows that the term $r_t - R_t + \theta_h$ should be a stationary cointegrating relation, *i.e.* in the long-run, the long-term rate should equal the short-term rate plus a constant term premium.
Therefore, a necessary condition for the EHTS to hold is that the long-term rate and the short-term rate are cointegrated, with cointegrating vector $(1,\ -1,\ \bar{\theta}_h)$, for every maturity h.

The second step is relevant and indispensable, because the first step only gives indications about "low frequency" implications of the theory. It consists in testing the cross-equation restrictions implied by the theoretical equation on the parameters of a stationary transformation of the statistical model as it will be shown in what follows. Therefore, in the empirical analysis, three hypotheses have to be verified, namely:

- the cointegration hypothesis (H_r), *i.e.* there must exist a cointegrating relation between interest rates at all maturities;

- the restrictions on the beta matrix (H_β), which maintain that the long-run cointegrating vectors are $(1,\ -1,\ \bar{\theta}_h)$, for every maturity h.

- the cross equation restrictions (H_{CER}) on the parameters of the VAR.

The three hypotheses have to hold jointly, in order for the EHTS to be valid. Any violation implies rejection of the theoretical model. As Bårdsen and Fanelli (2014) suggest, in testing many hypotheses using *e.g.* the 5% significance level for each hypothesis, the overall size of the test procedure is likely to be large (Savin, 1984). For this reason, the overall asymptotic size should not exceed the sum of the type I errors pre-fixed for each test.
As in Campbell and Shiller (1987) and Della Corte et al. (2008), the theoretical predictions will be tested on bivariate $VARs$ which model couples of short term rate and long term rates. Then, the testing procedure will be extended to a multivariate VAR process which jointly models the whole term structure of p maturities, in order to make a more reliable assumption on the Data Generating Process of the repo rates.

Hence, the interest rates of the term structure are assumed to follow a $VAR(k)$ process:

$$\mathbf{X}_t = \mathbf{c} + \mathbf{A}_1\mathbf{X}_{t-1} + ... + \mathbf{A}_k\mathbf{X}_{t-k} + \boldsymbol{\varepsilon}_t \tag{4}$$

with $\mathbf{X}_t$ being a $p \times 1$ vector. In the bivariate model, $\mathbf{X}_t := [r_t\ R_t^{(h)}]'$ and $p = 2$, whereas in the joint system $\mathbf{X}_t := [r_t\ R_t^{(1)}\ R_t^{(2)}...R_t^{(n)}]'$ is a $p = n+1$ vector of the short-term interest rate and the $R_t^{(h)}$, $h = 1, ..., n$, long-term interest rates.

Notice that equation (4) is a conditional model for $\mathbf{X}_t$ on its past values up to a truncation lag k, *i.e.* $E(\mathbf{X}_t \mid I_t) = E(\mathbf{X}_t \mid \mathbf{X}_{t-1}, \mathbf{X}_{t-2}, ..., \mathbf{X}_{t-k})$. Here, the information set is represented by past observations of interest rates. In this sense, this modeling assumption qualifies the approach as a "weak form" - type of Expectation Hypothesis testing, as defined by Fama (1970).

3.1 First step: rank test and restrictions on $\boldsymbol{\beta}$ (H_r and H_β)

As stated above, as a first implication of EHTS, the term structure of interest rates should be driven by a common stochastic trend and the $(p-1)$ spread(s) between long rate(s) and the short rate should be stationary. Consequently, the multivariate process (4) should be characterized by one common stochastic trend and $(p-1)$ cointegrating relations.
The long-run implications are the same both for the inexact present value model (*i.e.* equation (2) with $\tilde{\theta}_t^{(h)} \sim MDS$) and the exact present value model (*i.e.* equation (2) with $\tilde{\theta}_t^{(h)} = 0$) as remarked by González and Gonzalo (2000).
Considering the $VECM$ representation of the VAR in (4)

$$\Delta\mathbf{X}_t = \boldsymbol{\Pi}\mathbf{X}_{t-1} + \sum_{i=1}^{k-1} \boldsymbol{\Gamma}_i \Delta\mathbf{X}_{t-i} + \boldsymbol{\varepsilon}_t \tag{5}$$

the aforementioned condition translates into a reduced rank equal to $(p-1)$ of the matrix $\mathbf{\Pi}$.
In order to statistically test this prediction, rank tests are applied, namely, Johansen's test, which relies on asymptotic distribution of the test statistics, and Cavaliere et al. (2012, 2014) test, which tackles the problem of heteroskedasticity with wild bootstrap.
If the condition is satisfied, equation (5) can be written as:

$$\Delta \mathbf{X}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{X}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{X}_{t-i} + \boldsymbol{\varepsilon}_t \tag{6}$$

with $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ matrices describing the short-run and the long-run dynamics, respectively. In the bivariate case, the long-run matrix $\boldsymbol{\beta}$ has dimension 2×1, whereas in the model for all interest rates $\boldsymbol{\beta}$ is a $p \times n$ matrix.
In the bivariate model, the cointegrating vector should have as coefficients $(1,\ -1,\ \bar{\theta}_h)$.
As for the joint system, the columns of the $\boldsymbol{\beta}$ which outline the cointegrating relations should be as such:

$$\boldsymbol{\beta} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \\ \bar{\theta}_1 & \bar{\theta}_2 & \cdots & \bar{\theta}_n \end{bmatrix} \tag{7}$$

This can be tested by leaving the first row of parameters unrestricted and then verifying the restrictions with a Likelihood Ratio (LR) test as in Hansen (2003).
Again, together with the asymptotic test, a bootstrap test (Boswijk et al., 2013) is performed in the testing procedure.
Of particular relevance is the fact that, both in the bivariate and in the joint system, the $VECM$ allows to identify the n long-run constant component of the term premia $\bar{\theta}_h,\ h = 1, .., n$ in the matrix $\boldsymbol{\beta}$. [5]

3.2 Second step: cross equation restrictions (H_{CER})

Following Campbell and Shiller (1987), once the necessary condition has been fulfilled, the restrictions implied by the rational expectation model in equation (3) should be tested in a stationary VAR model.
For the derivation of the constraints implied by a bivariate VAR the reader can refer to previous literature and in particular to González and Gonzalo (2000) for the inexact present value model's restrictions.
In what follows, the derivation of the constraints implied from the system of all interest rates is discussed.

[5] Term premia could be modeled as piece-wise time varying in a $VECM$ model with structural changes as proposed by Hansen (2003).

The $VECM$ model in equation (5) is reparametrized as the following stationary VAR:

$$\mathbf{Y}_t = \mathbf{B}_1 \mathbf{Y}_{t-1} + ... + \tilde{\mathbf{B}}_k \mathbf{Y}_{t-k} + \mathbf{e}_t \tag{8}$$

where

$$\mathbf{Y}_t := \begin{bmatrix} \check{\mathbf{S}}_t \\ \Delta r_t \end{bmatrix}$$

with $\check{\mathbf{S}}_t$ being an $(n \times 1)$ vector of cointegrating relations, *i.e.* the spreads minus the respective estimated long-run constant term premium $\hat{\bar{\theta}}_h$, $\check{S}_t^{(h)} := R_t^{(h)} - r_t - \hat{\bar{\theta}}_h$, $h = 1, ..., n$. The last column of $\tilde{\mathbf{B}}_k$ is restricted to be $\mathbf{0}_p$.
The companion form of equation (8) is

$$\mathbf{Y}_t^* = \mathbf{J}_B \ \mathbf{Y}_{t-1}^* + \mathbf{e}_t^*$$

The theoretical equations (2) for each maturity can put together in a system which, in compact form, would be:

$$\check{\mathbf{S}}_t = \mathbf{M}_\delta E_t \check{\mathbf{S}}_{t+1} + \mathbf{D}_\delta E_t \Delta r_{t+1} + \tilde{\boldsymbol{\theta}}_t \tag{9}$$

with $\tilde{\boldsymbol{\theta}}_t$ being either equal to zero (exact present value model) or an MDS (inexact present value model).
The variables of equation (9) are expressed in terms of $\mathbf{Y}_t^*$, so that, through substitutions, it is possible to obtain the following expression:

$$\mathbf{R}_S - \mathbf{M}_\delta \mathbf{R}_S \mathbf{J}_B - \mathbf{D}_\delta \mathbf{R}_r \mathbf{J}_B = \mathbf{0}$$

for the exact present value model, and

$$\mathbf{R}_S \mathbf{J}_B - \mathbf{M}_\delta \mathbf{R}_S \mathbf{J}_B^2 - \mathbf{D}_\delta \mathbf{R}_r \mathbf{J}_B^2 = \mathbf{0}$$

in the case of inexact formulation.
The restrictions on matrix $\mathbf{J}_B$ (hence on the parameters of the equation (8)), are derived from the above equations.
The example in Appendix A clarifies how the restrictions are formulated.

The restrictions derived as such, are tested on the stationary VAR of equation (8). In the model for all interest rates, the joint restrictions are very likely to be rejected because of the complexity of the constraints. Instead, the hypothesis tests of the n EHTS equations (3) verified on bivariate VARs, breaks down the complexity of the problem.
In order to treat heteroskedasticity, the bootstrap approach proposed by Hafner and Herwartz (2009) in a stationary VAR framework and HAC correction are adopted.

3.3 Rolling window analysis

With the aim of taking into account a more general type of parameters variation over time, a rolling window analysis is performed. In this case, not only the variance-covariance matrix, but also all the parameters of the statistical model are assumed to change over the period under investigation. In this way, it is also possible to follow the evolution of the test statistic over time and evaluate in which periods the EHTS holds.

The rolling window analysis consists in computing the test statistics for the hypotheses H_r, H_β and H_{CER} and the relative p-values for each sub-sample (window) of 1000 observations (approximately 4 years with business days data), *i.e.* for sub-samples (1,...,1000), (2,...,1001), ... , (4576,..., 5576).

Since H_r, H_β and H_{CER} are tested jointly, each sub-sample will not reject the EHTS when the sum of the p-values of all the tests will exceed the pre-fixed level of significance. Potentially, the critical values of each test could be α/N where N is the number of hypothesis to be tested jointly.

4 Data

The methodology described above is applied to a dataset of general collateral government repo rates. The data consists of daily observations of closing repo rates for the following maturities: overnight, one-week, two-weeks, three-weeks, one-month, two-months, and three-months, so that the term structure comprises 7 interest rates. The sample goes from May 21^{st} 1991 to October 23^{rd} 2013.

For comparative purposes, the analysis is also conducted in a sample which excludes the observation from May 1^{st} 2009[6] on - in order to exclude the too stable last part of the sample - and on Della Corte et al. (2008) sample (May 21^{st} 1991 - December 9^{th} 2005).

An extensive descriptive analysis of the data set has been provided by Longstaff (2000) and Della Corte et al. (2008) for the sample May 21^{st} 1991 - December 9^{th} 2005. As it is possible to see in Figure 1, the most relevant aspect of the following period included in the present analysis, is the sharp increase of repo rates until the recent financial crisis. After the burst of the crisis, the levels of variables decrease and stabilize around the zero bound.[7]

The time series plot of the levels and first differences of the interest rates (Figure 1 and Figure 2, respectively) show two relevant aspect for the present analysis: non-stationarity and heteroskedasticity.

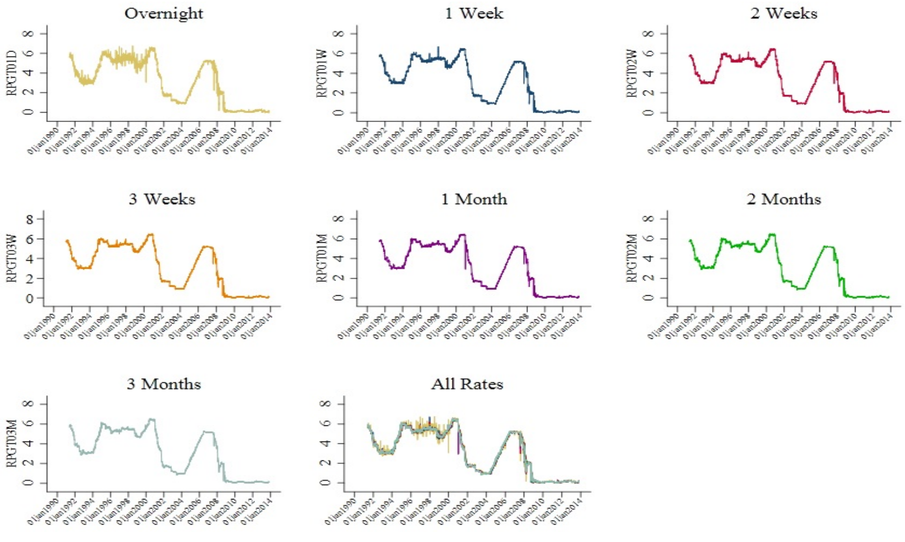
For what regards the former, the series show clear signs of non-stationarity and the statistical analysis following in the next section confirms this finding.

Moreover, periods of high volatility and low volatility are alternating all over the sample. In particular, higher noise coincides with financial markets' crises. Furthermore, shorter maturities display higher volatility.

[6]Since 2009, the Fed undertook "unconventional monetary policies" which had the effect of stabilizing the interest rates to a low level. Furthermore, since 1^{st} May 2009 the fail charge was introduced by the Fed (see `http://www.newyorkfed.org/tmpg/faq.html`), *i.e.* up to 3% charge in case of fail to deliver upon trade.

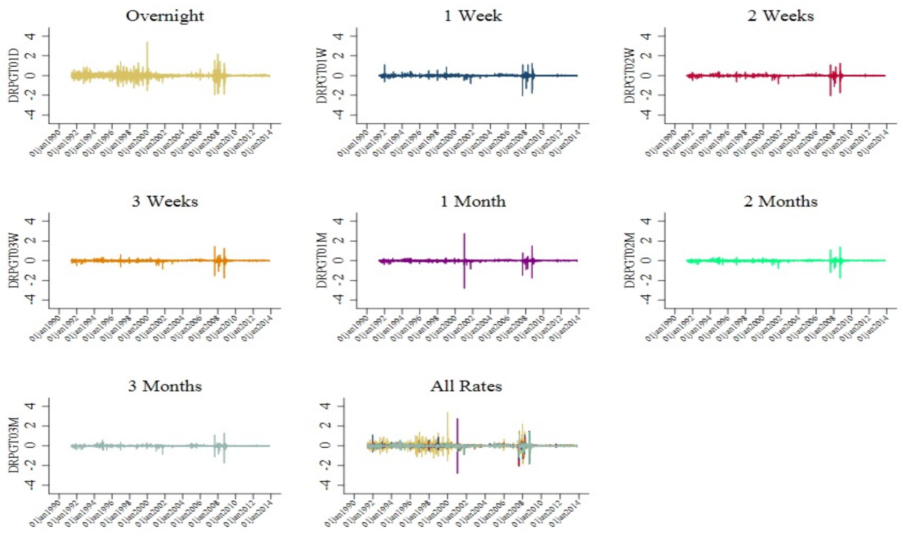
[7]General Collateral Repo are backed by very liquid assets (generally Treasury bills maturing in less than 10 years), differently from Special Repo contracts, in which a specific asset is asked for as collateral. Therefore, the interest rate of the former is that one prevailing in the money market and it is really close to rates on overnight loans in the federal funds market. It follows that repo rates are closely related to the FED rate, which was brought to near zero in response to the financial crisis.

Figure 1: Repo interest rates



Notes: The graphs show time series plots of the interest rates at daily frequency (business days). Sample: 21^{st} 1991 - October 23^{rd} 2013.

Figure 2: Repo interest rates - first differences



Notes: The graphs show time series plots of first differences of interest rates at daily frequency (business days). Sample: 22^{st} 1991 - October 23^{rd} 2013.

5 Results

The implementation of the two steps begins with the estimation of VAR models for the interest rates at different maturities.The VAR's lag length is fixed at $k = 2, 5$ and 11, as suggested by previous literature and by lag selection information criteria[8]. However, results are invariant to the sample choice. In what follows, results for lag $k = 5$ and May 21^{st} 1991 - October 23^{rd} 2013 sample are presented. Results for other samples can be found in the Appendix B.

5.1 First step: rank test and restrictions on $\boldsymbol{\beta}$ (H_r and H_β)

The first step is based on testing the long-run implications of the EHTS, *i.e.* the presence of a common stochastic trend driving the rates of the term structure. This is empirically translated the $\boldsymbol{\beta}$ matrix in equation (6) having column rank equal to $n = 1$ in the bivariate system case and column rank equal to $n = 6$ in the joint system case, *i.e.* in the rank hypothesis (H_r) testing.
The tests' results are shown in Table 1.
Johansen (1995) rank test and the heteroskedastic variance Cavaliere et al. (2012, 2014) test[9] do not reject the null of cointegration rank equal to $(n - 1)$, both in the bivariate (Table 1 - Panel A) and in the multivariate case (Table 1 - Panel B). Fixed the cointegration rank, the $VAR(5)$ can be expressed in its $VECM$ form (equation (6)) and the cointegrating restrictions (7) implied by the theoretical model can be tested on the $\boldsymbol{\beta}$ matrix (H_β).
The results of the LR test for the h bivariate systems are shown in Panel A of Table 2, together with the estimated constant term premia $\bar{\theta}_h$. Both the asymptotic test and the bootstrap test by Boswijk et al. (2013) cannot reject the null hypothesis.
Panel B of Table 2 reports the LR tests and the $\bar{\theta}_h$ for the joint system test. Also in this case, the test on the $\boldsymbol{\beta}$ matrix cannot reject the restrictions implied by the EHTS, both in the asymptotic version and in the bootstrap version (see Panel A).
The long-run constant term premia $\hat{\bar{\theta}}_h$ $h = 1W, 2W, 3W, 1M, 2M, 3M$ are estimated and reported in the first column of both Panels.

The estimation of the constant term premia from the multivariate model should be more reliable, because the information set on which the model is based is richer. The estimated term premia of the model for all interest rates tend to be significant but very small in magnitude (at most 0.0399 basis points). Agents ask for a very small extra-yield, given the fact that, when choosing longer maturities, they have to commit for a small additional amount of time with respect to the overnight maturity. Moreover, the term premia are increasing with maturity, as expected. Surprisingly, the term premia related to the 1W and the 2W rates are negative. A tentative explanation for the negative signs is that the cost of committing for a

[8]Both Longstaff (2000) and Della Corte et al. (2008) use 5 lags. For what regards information criteria, the Schwartz Bayesian Criterion (SBC) chooses 2 lags, Hannan and Quinn Criterion (HQC) 11 lags and Akaike Criteria (AIC) 12, using a maximum lag equal to 12. The choice of a bigger maximum lag would have led to the selection of a higher lag for the AIC, but Ljung-Box Q test of no-autocorrelation shows total absence of residuals' autocorrelation already in a model with $k = 6$.

[9]The boostrap testing procedure has been implemented with the command **bootrank.ado** for Stata which I programmed on purpose. However, the ado file is meant to be used on every dataset.

Table 1: **Cointegration Rank Test -** H_r

Deterministic component: restricted constant	obs =	5576
Sample: May 29^{th} 1991- October 23^{rd} 2013	Lags =	5

Panel A: Bivariate systems

	rank	**trace stat.**	**P-value**	**bootstrap p-value**
1W	0	854.29	0.000	0.000
	1	2.44	0.691	0.671
2W	0	730.75	0.000	0.000
	1	2.58	0.665	0.619
3W	0	624.72	0.000	0.000
	1	2.84	0.616	0.553
1M	0	535.17	0.000	0.000
	1	2.84	0.617	0.607
2M	0	421.50	0.000	0.000
	1	3.21	0.550	0.551
3M	0	334.06	0.000	0.000
	1	3.01	0.586	0.581

Panel B: Joint systems

rank	**trace stat.**	**P-value**	**bootstrap p-value**
0	4492.70	0.000	0.000
1	3278.22	0.000	0.000
2	2286.93	0.000	0.000
3	1450.74	0.000	0.000
4	802.72	0.000	0.000
5	220.02	0.000	0.000
6	2.33	0.712	0.676

Notes: In the last two columns p-values of Johansen (1995) rank test and Cavaliere et al. (2012, 2014) bootstrap rank test are reported. For the latter, Gaussian wild bootstrap has been performed. The number of bootstrap replications is 499.

longer period is lower that the cost of rolling over the one period repo contract, since the commitment is for such a short time period. Therefore, the agents would prefer to renounce to some yield in order to secure his position for one or two weeks, hence the negative term-premia.

Table 2: **Test on $\boldsymbol{\beta}$ - H_β**

	Panel A: Bivariate systems				Panel B: Joint system			
	$\bar{\theta}_h$	**LR stat.**	**P-value**	**bootstrap p-value**	$\bar{\theta}_h$	**LR stat.**	**P-value**	**bootstrap p-value**
1W	-0.0059 (0.002)	5.27	0.022	0.033	-0.0052 (0.000)			
2W	-0.0039 (0.003)	2.47	0.116	0.097	-0.0025 (0.000)			
3W	-0.0014 (0.004)	0.95	0.329	0.294	0.0006 (0.000)	9.43	0.151	0.196
1M	0.0053 (0.005)	0.19	0.666	0.635	0.0085 (0.000)			
2M	0.0184 (0.007)	0.18	0.671	0.709	0.0229 (0.000)			
3M	0.0345 (0.009)	0.82	0.36	0.371	0.0399 (0.000)			

Notes: LR test compares the restricted model with $\boldsymbol{\beta}$ matrix as (7) with the unrestricted model that has a free parameter on the first row of (7) in the $VECM$ equation $\Delta\mathbf{X}_t = \mathbf{\Pi}\mathbf{X}_{t-1} + \sum_{i=1}^{4}\mathbf{\Gamma}_i\Delta\mathbf{X}_{t-i} + \boldsymbol{\varepsilon}_t$. Panel A and Panel B report the results for the test on the bivariate system and the joint system, respectively. In each Panel, the first column reports the estimates of the constant term premia $\bar{\theta}$ (standard errors in brackets). Last two columns display p-values of Johansen (1995) rank test and Boswijk et al. (2013) bootstrap LR test.

5.2 Second step: cross equation restrictions (H_{CER})

As mentioned above, the second step verifies the cross-equation restrictions implied by the theoretical model on the VAR of equation (8). In what follows, constraints from both the exact present-value model and the inexact present value model are tested in a VAR framework.
In the case of the bivariate model, with $k = 5$, the restrictions derived from the VAR ammount to 9.
The constraints implied by the EHTS on the VAR model comprising all the interest rates are derived as explained in Appendix A. The restrictions amount to a very large number, which is increasing with the number of lags considered[10].
In the joint system case, the set of restrictions is rather complex, hence the null hypothesis is very likely to be rejected. The rejection of the EHTS is not the only reason for this outcome.
Table 3 shows the results of the cross-equation restriction tests for the n bivariate VARs. The two Wald-type statistics proposed by Hafner and Herwartz (2009) are computed in order to adjust for the effect of heteroskedasticity. Table 3 shows, for each long-term maturity, the test statistics and the p-values associated with the test. In the case of non-linear restrictions (inexact present value model), it is not possible to apply the bootstrap procedure suggested by the authors because a closed-form solution for the coefficients under the non-linear constraint does not exist, hence the

[10]For instance, for a $VAR(5)$ the restrictions to be imposed are $n(nk) - n = 204$ (see Appendix A for details).

bootstrap samples cannot be generated under the null hypothesis.
As it is possible to see, H_{CER} is strongly rejected at any significance level for almost all the maturities in the case of the exact present value model, exceptions being the 2 weeks, 3 weeks and 1 month rates (with heteroskedasticity correction, either with the bootstrap procedure or the White approach).
However, since the EHTS is tested for the whole term structure in a multiple hypothesis testing framework, H^h_{CER} should not be rejected at any maturity, given the pre-fixed significance level. This is the case of the test statistics corrected for heteroskedasticity for the inexact present value model.

Table 3: **Test of restrictions implied by EHTS H_{CER} - bivariate $VARs$**

	Exact Model		**Inexact Model**	
	Wald test	**Wald test with HAC**	**Wald test**	**Wald test with HAC**
1W	345.51 (0.000) (*0.012*)	31.16 (0.000) (*0.000*)	173.27 (0.000)	13.67 (0.135)
2W	225.40 (0.000) (*0.166*)	31.82 (0.000) (*0.000*)	71.89 (0.000)	8.35 (0.500)
3W	339.03 (0.000) (*0.040*)	22.57 (0.007) (*0.002*)	73.76 (0.000)	6.03 (0.737)
1M	875.26 (0.000) (*0.098*)	11.08 (0.268) (*0.251*)	22.16 (0.008)	3.92 (0.917)
2M	545.74 (0.000) (*0.004*)	57.57 (0.000) (*0.000*)	32.09 (0.000)	7.15 (0.622)
3M	600.42 (0.000) (*0.000*)	80.37 (0.000) (*0.000*)	54.69 (0.000)	15.69 (0.074)

Notes: The Wald tests (plain and with White (1980) Heteroskedasticity and Autocorrelation Consistent (HAC) covariance, see Hafner and Herwartz, 2009) verify the restrictions implied by the theoretical equation $\breve{S}^{(h)}_t = \delta_h E_t \breve{S}^{(h)}_{t+1} + \delta_h E_t \Delta r_{t+1} + \tilde{\theta}^{(h)}_t$ with $\tilde{\theta}^{(h)}_t = 0$ (*i.e.* exact present value model) or $\tilde{\theta}^{(h)}_t \sim MDS(\mathbf{0}, \mathbf{\Sigma}_\theta)$ (*i.e.* "inexact" present value model) on the VAR model $\mathbf{Y}_t = \mathbf{B}_1\mathbf{Y}_{t-1} + ... + \tilde{\mathbf{B}}_5\mathbf{Y}_{t-5} + \mathbf{e}_t$, with $\mathbf{Y}_t = [\breve{S}^{(h)}_t, \Delta r_t]'$ and $h = 1W, 2W, 3W, 1M, 2M, 3M$. In brackets, asymptotic and bootstrap (in italics) p-values are reported. Bootstrap p-values are computed with $B = 499$ replications.

Table 4 reports the result of the Likelihood Ratio (LR) test of H_{CER} on the VAR including all interest rates. Given the high number of restrictions (204) implied by the VAR, a Wald-type test is algebraically too involving to be formulated.

Table 4: **Test of restrictions implied by EHTS H_{CER} - joint system**

Trend: restricted constant	obs =	5576
Sample: May 29^{th} 1991- October 23^{rd} 2013	Lags =	5
	LR(204)	p-value
exact	14115.69	0.000
inexact	11529.89	0.000

Notes: The LR test verifies the restrictions implied by the theoretical equation $\check{\mathbf{S}}_t = \mathbf{M}_\delta E_t \check{\mathbf{S}}_{t+1} + \mathbf{D}_\delta \Delta r_{t+1} + \tilde{\theta}_t$ with $\tilde{\theta}_t = 0$ (*i.e.* exact present value model) or $\tilde{\theta}_t \sim MDS(\mathbf{0}, \mathbf{\Sigma}_\theta)$ (*i.e.* "inexact" present value model) on the VAR model $\mathbf{Y}_t = \mathbf{B}_1 \mathbf{Y}_{t-1} + ... + \tilde{\mathbf{B}}_5 \mathbf{Y}_{t-5} + \mathbf{e}_t$, with $\mathbf{Y}_t = [\check{S}_t^{(1W)}, ..., \check{S}_t^{(3M)}, \Delta r_t]'$. The p-value is reported in the last column.

As perhaps expected, the restrictions are strongly rejected. This might be due to the complexity of the restrictions involved.[11]
An alternative method for testing the restrictions implied by the EHTS in a joint system, without explicitly deriving the restrictions, could be the "graphical method". The methodology was introduced in Campbell and Shiller (1987) and revised by Johansen and Swensen (2011) by using parameters from the $VECM$ model in equation (5). The method consists in building the spread series as predicted by the theoretical model and compare them with the actual spread series. This can be object of future research.

5.3 Multiple hypotheses testing

Summing up the results for the whole testing procedure, in the bivariate testing case, if an overall significance level of 5% and hence a type-I error of the tests equal to 0.00625 is considered, each stage of the procedure (*i.e.* H_r, H_β and H_{CER}) is not rejected[12] in the inexact present value model case, as shown in the summarizing table 7.
Notice that the non-rejection of the EHTS is verified once time varying volatility is taken into account in the statistical model. This result underlines the importance of heteroskedasticity correction in Present Value models.
Overall, the empirical analysis shows supporting evidences in favor of the EHTS, in the model which allows for transitory deviations from the equilibrium conditions, *i.e.* in the inexact present value model.

[11] Bootstrap inference appears to be useless, because it will not lead to the non-rejection of the null hypothesis. Moreover, a bootstrap procedure has been proposed by Hafner and Herwartz (2009) for a Wald-type test, which is not possible to compute for the reasons explained above.

[12] The reader should remind that under the null the constraints implied by the EHTS are true.

Table 5: **Test of Expectation Hypothesis - Summary results H_r, H_β, H_{CER}**

		Rank test - H_r			**Beta test - H_β**			**CER test - H_{CER}**				
								Exact model			**Non exact model**	
		trace stat.	**p-value**	**bootstrap p-value**	**LR stat.**	**p-value**	**bootstrap p-value**	**Wald(HAC) stat**	**p-value**	**bootstrap p-value**	**Wald(HAC) stat**	**p-value**
Bivariate System	*1W*	2.44	0.691	0.671	5.27	0.022	0.033	31.16	0.000	0.000	13.67	0.135
	2W	2.58	0.665	0.619	2.47	0.116	0.097	31.82	0.000	0.000	8.35	0.500
	3W	2.84	0.616	0.553	0.95	0.329	0.294	22.57	0.007	0.002	6.03	0.737
	1M	2.84	0.617	0.607	0.19	0.666	0.635	11.08	0.268	0.251	3.92	0.917
	2M	3.21	0.550	0.551	0.18	0.671	0.709	57.57	0.000	0.000	7.15	0.622
	3M	3.01	0.586	0.581	0.82	0.364	0.371	80.37	0.000	0.000	15.69	0.074
Joint System		2.33	0.712	0.676	9.43	0.151	0.196	14115.69	0.000		11529.89	0.000

Notes: The table reports the results of the rank test (H_r), the test on the beta coefficients (H_β) and the test of the cross-equation restrictions (H_{CER}) from the previous tables 1-3 (see for further details). Asymptotic and bootstrap p-values are reported.

5.4 Rolling window analysis

The graphs in Figure 3 and 4 show the rolling window of the p-values for the testing of the rank hypothesis (H_r), the test of restrictions on the beta matrix (H_β) and the test of CER for each maturity (H_{CER}) for the exact and the inexact bivariate models, respectively[13]. The EHTS is rejected when one of the test rejects its null hypothesis[14].

In the graphs, each test rejects its null hypothesis (*i.e.* the validity of EHTS) when the p-value series is below the significance level (red line). The latter is pre-fixed at $0.05/3 = 0.01666$ because of multiple hypothesis testing.The grey areas highlight the rejection periods.

In the exact model, the EHTS is rejected most of the times, the grey area covering almost the whole graph area.

As for the inexact model, the rolling window analysis shows that the long-run implication of EHTS are confirmed by data most of the time. Considering each maturity, it is possible to notice that the shortest maturities reject the EHTS more frequently along the sample.

However, in some periods the EHTS is definitely rejected for all maturities. Interestingly, these periods coincide with times of turmoil of the financial market and/or the repo market. After all, it is reasonable to conjecture that the rational expectation and unlimited arbitrage hypothesis which are on the basis of the EHTS do not hold during economic crises[15].

The first rejections are between the Asian and the Russian financial crises and in the aftermath of the Russian crisis (late 1998-1999). Likewise, the Dot-com bubble burst between 2000 and 2001 and the 9/11 caused disruption in the financial as well as in the repo market in particular. Because of the terrorist attack, settlement fails had an enormous jump (more than 110%).[16] For what regards the years between 2003 to 2006, this period was characterized by many fails on repo contracts[17] and by "market squeezes" which distorted prices in the repo market (see Remarks of Deputy Assistant Secretary for Federal Finance James Clouse U.S. Department of the Treasury, September 27^{th}, 2006). As for the first issue, in April 2006 the Bond Market Association adopted the fail penalty measure, which charges a penalty on the trader in case of fail. For what regards market squeezes, some big traders - deemed to manipulate Treasury and repo market - were fired in the last months of 2006 and no market squeeze materialized anymore. As it is possible to see on the graph, after these events, the EHTS equilibrium was restored until the recent financial crisis. In this period, the repo market was affected by numerous fails and

[13] In the joint system, the cross-equation restrictions (H_{CER}) are rejected all over the sample. Therefore the rolling windows graph is not reported.

[14] In some windows, the maximization of the log-likelihood function was not possible because of collinearity. Therefore, the LR statistics was not computable and the corresponding observations have been dropped.

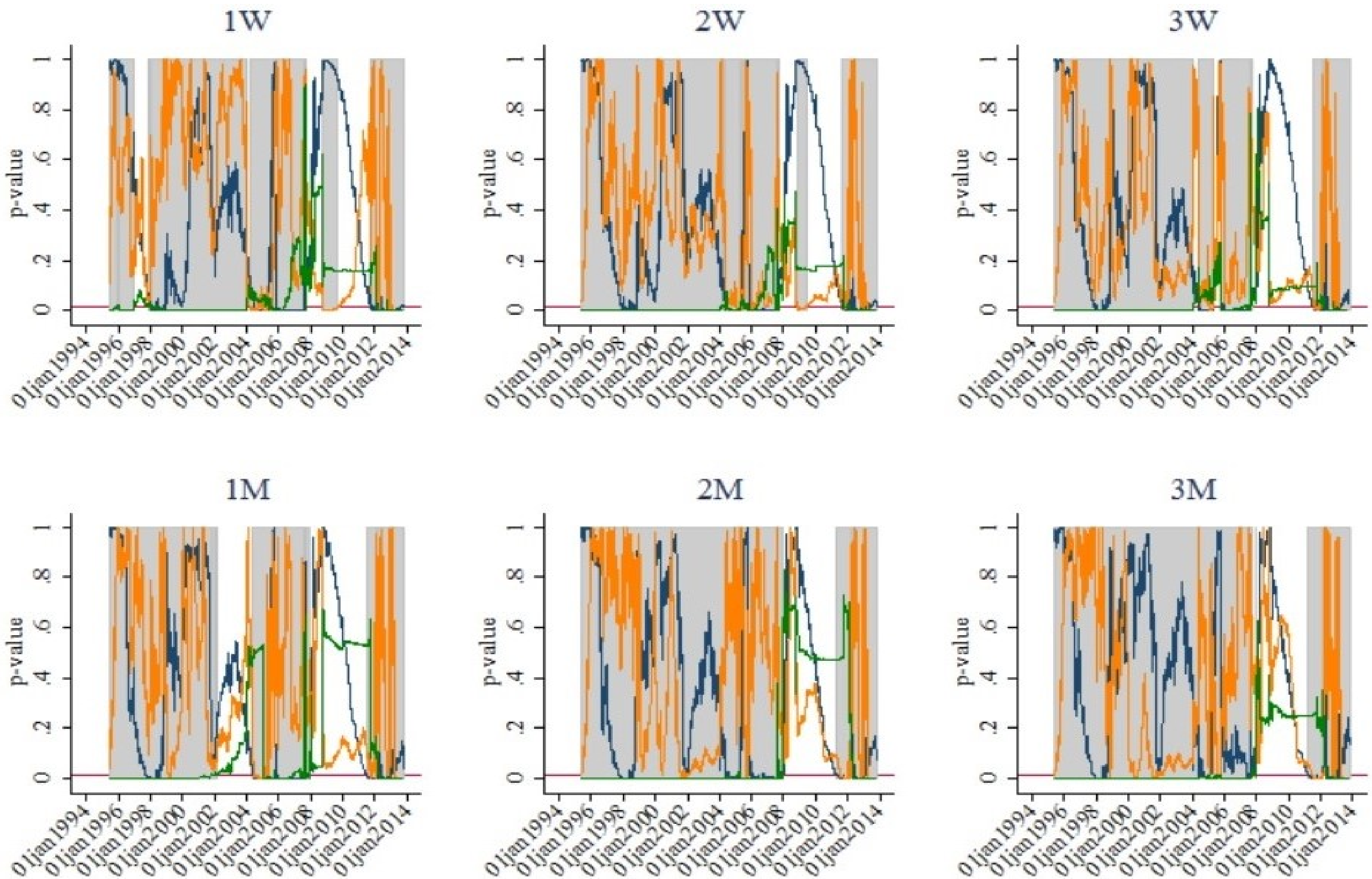
[15] The reader can refer to Krishnamurthy (2009) for an account of limit of arbitrage problems in the repo market during the recent crisis.

[16] See Fleming and Garbade (2002).

[17] Even though counterparty fail of a repo contract has no nominal costs - in such an event, the counterpart can keep the collateral - it implies back-office costs and negative externalities on traders' business.

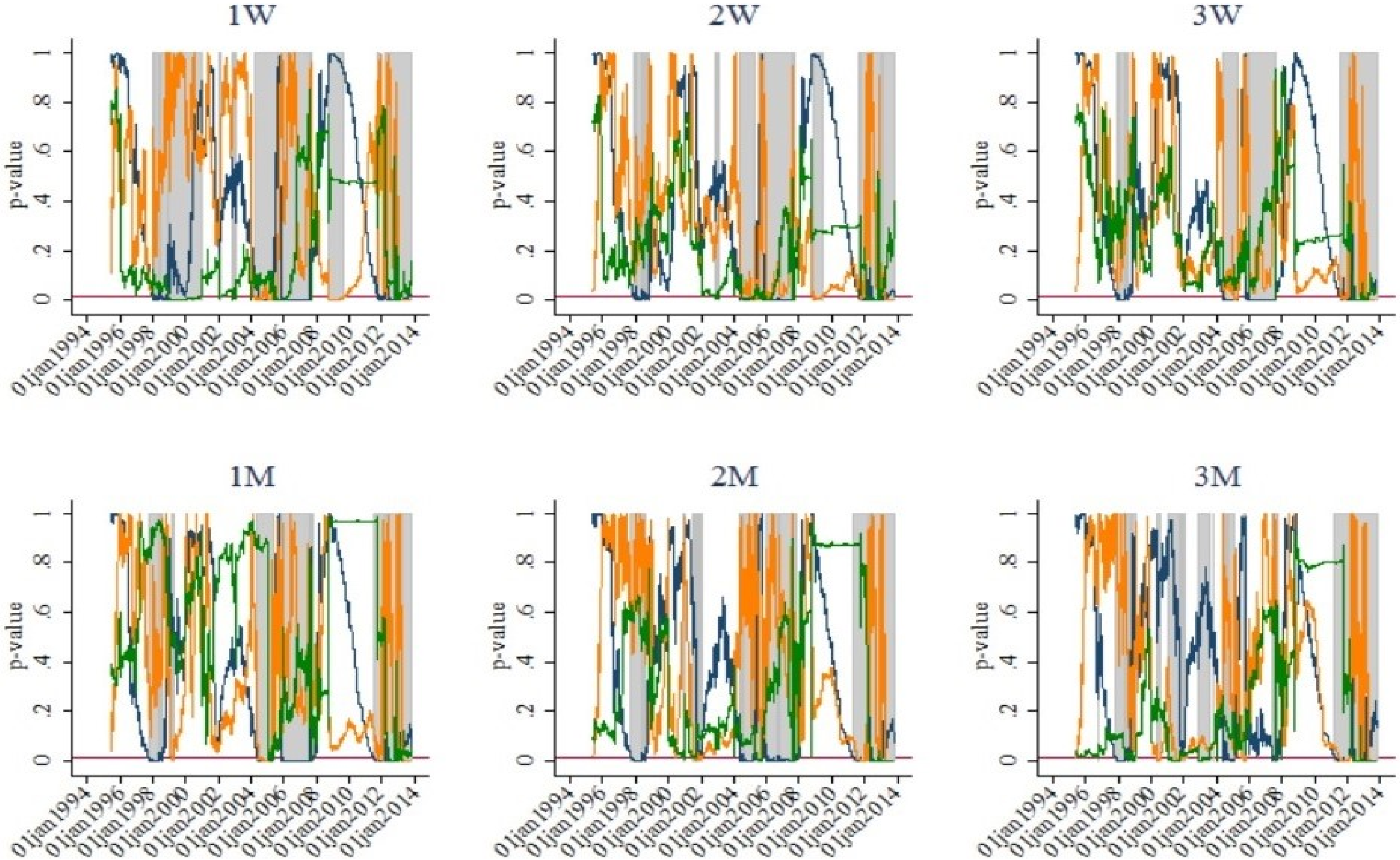
run on repo contracts, especially those backed by private sector collaterals (see Gorton and Metrick, 2012 and Krishnamurthy et al., 2012). The steeper penalty fail introduced by the Fed on May 1^{st} 2009, mitigated the repo market distress. Right after this event, the EHTS test starts to not-reject again the null, until July 2011, when concerns about US debit ceiling started to spread among financial sector's operators (see D. Duffie and A K Kashyap, "US default would spell turmoil for the repo market", The Financial Times, July 29^{th} 2011).

Figure 3: Rolling window p-values - w.s.=1000 - Exact Model



Notes: The Figure reports the graphs of the rolling window p-values for each of the test, in the following order: rank test (blue line), test of restrictions on the beta matrix (orange line) and the CER (green line) for each maturity ("1W",...,"3M"). For each date, the corresponding p-value of the relevant test are reported. The red line represents 0.01666 significance level. A window size of 1000 observation has been used.

Figure 4: Rolling window p-values - w.s.=1000 - Inexact Model



Notes: The Figure reports the graphs of the rolling window p-values for each of the test, in the following order: rank test (blue line), test of restrictions on the beta matrix (orange line) and the CER (green line) for each maturity ("1W",...,"3M"). For each date, the corresponding p-value of the relevant test are reported. The red line represents 0.01666 significance level. A window size of 1000 observation has been used.

6 Conclusions

The empirical verification of the EHTS has always constituted a controversial ground for applied econometrics research. Many empirical strategies have been applied, with mixed conclusions.
For what regards the repo market, two previous studies by Longstaff (2000) and Della Corte et al. (2008) provide contrasting statistical evidences by using different methodologies.

This paper applies most recent econometric techniques to provide a rigorous statistical treatment of the testing problem. The approach adopted allows to control for the following features of the data and of the testing procedure in a coherent framework:

1. the possibility of temporaty deviation of the data from the equilibrium which determines the EHTS equation ("inexact" form);
2. the non-stationarity of the repo interest rates;
3. the heteroskedasticity in the data;
4. the joint nature of the statistical problem;
5. the time-dependency of the validity of the EHTS.

The empirical results provide overall support to the validity of EHTS. None of the tests rejects the restrictions implied by the EHTS in the inexact model. It is important to notice that the result is reached only after heteroskedasticity correction.
Finally, a rolling window analysis shows that the implications of the EHTS are only rejected in periods of turmoils either of the financial or of the repo market.

Appendices

A Restrictions implied by the EHTS - Example: $n = 2$, $k = 2$

Consider the case in which there are a short-run interest rate and two long-run interest rates. In this case, the dynamics of the term structure of interest rates is modeled as the following multivariate process:

$$\mathbf{X}_t := \begin{bmatrix} r_t \\ R_t^{(1)} \\ R_t^{(2)} \end{bmatrix}$$

As outlined above, the tri-variate system should have one common trend and two cointegrating relations which are the spreads (to which the respective estimate of the constant term premium is subtracted) $\check{S}_t^{(1)} := R_t^{(1)} - r_t - \hat{\bar{\theta}}_1$ and $\check{S}_t^{(2)} := R_t^{(2)} - r_t - \hat{\bar{\theta}}_2$. Moreover, the vector $\mathbf{Y}_t$ of equation (8) is defined as such:

$$\mathbf{Y}_t := \begin{bmatrix} \check{\mathbf{S}}_t \\ \Delta r_t \end{bmatrix} := \begin{bmatrix} \check{S}_t^{(1)} \\ \check{S}_t^{(2)} \\ \Delta r_t \end{bmatrix}$$

The theoretical relation to be tested (equation 3) is a bivariate system:

$$\begin{bmatrix} \check{S}_t^{(1)} \\ \check{S}_t^{(2)} \end{bmatrix} = \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix} \begin{bmatrix} E_t \check{S}_{t+1}^{(1)} \\ E_t \check{S}_{t+1}^{(2)} \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} E_t \Delta r_{t+1} + \begin{bmatrix} \tilde{\theta}_t^{(1)} \\ \tilde{\theta}_t^{(2)} \end{bmatrix}$$

which can be rewritten in a compact form as:

$$\check{\mathbf{S}}_t = \mathbf{M}_\delta E_t \check{\mathbf{S}}_{t+1} + \mathbf{D}_\delta E_t \Delta r_{t+1} + \tilde{\boldsymbol{\theta}}_t \tag{10}$$

A.1 Exact present value model - $\tilde{\boldsymbol{\theta}}_t = 0$

In the case of a model which excludes time varying risk premia, the theoretical equation to be tested is:

$$\mathbf{S}_t = \mathbf{M}_\delta E_t \mathbf{S}_{t+1} + \mathbf{D}_\delta \Delta r_{t+1} \tag{11}$$

As illustrative example, consider the case in which $k = 2$. The VAR model in equation (8) is:

$$\mathbf{Y}_t = \mathbf{B}_1 \mathbf{Y}_{t-1} + \tilde{\mathbf{B}}_2 \mathbf{Y}_{t-2} + \mathbf{e}_t$$

and the companion form is:

$$\begin{bmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1 & \tilde{\mathbf{B}}_2 \\ \mathbf{I}_3 & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{t-1} \\ \mathbf{Y}_{t-2} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_t \\ \mathbf{0} \end{bmatrix}$$

$$\underset{6\times1}{\mathbf{Y}_t^*} = \underset{6\times6}{\mathbf{J}_B}\ \underset{6\times1}{\mathbf{Y}_{t-1}^*} + \underset{6\times1}{\mathbf{e}_t^*} \tag{12}$$

The restrictions implied by the rational expectation equation (10) will be expressed as restrictions on the parameters of the $\mathbf{J}_B$ matrix and ultimately on those of $\mathbf{B}_1$ and $\tilde{\mathbf{B}}_2$.
As it is possible to notice, each variable of equation (10) can be expressed in terms of $\mathbf{Y}_t^*$ as follows:

$$\mathbf{S}_t = \mathbf{R}_S\mathbf{Y}_t^*$$

$$E_t\mathbf{S}_{t+1} = \mathbf{R}_S E_t\mathbf{Y}_{t+1}^* = \mathbf{R}_S\mathbf{J}_B\mathbf{Y}_t^*$$

$$E_t\Delta r_{t+1} = \mathbf{R}_r E_t\mathbf{Y}_{t+1}^* = \mathbf{R}_r\mathbf{J}_B\mathbf{Y}_t^*$$

where

$$\mathbf{R}_S := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{R}_r := [0\ 0\ 1\ 0\ 0\ 0]$$

The substitution of the previous expressions in (10), leads to the following equation:

$$\mathbf{R}_S\mathbf{Y}_t^* = \mathbf{M}_\delta\mathbf{R}_S\mathbf{J}_B\mathbf{Y}_t^* + \mathbf{D}_\delta\mathbf{R}_r\mathbf{J}_B\mathbf{Y}_t^*$$

$$(\mathbf{R}_S - \mathbf{M}_\delta\mathbf{R}_S\mathbf{J}_B - \mathbf{D}_\delta\mathbf{R}_r\mathbf{J}_B)\mathbf{Y}_t^* = \mathbf{0}$$

since $\mathbf{Y}_t^* \neq \mathbf{0}$ a.s., the following equation must hold:

$$\mathbf{R}_S - \mathbf{M}_\delta\mathbf{R}_S\mathbf{J}_B - \mathbf{D}_\delta\mathbf{R}_r\mathbf{J}_B = \mathbf{0}$$

that is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{J}_B - \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} [0\ 0\ 1\ 0\ 0\ 0]\mathbf{J}_B = \underset{2\times6}{\mathbf{0}}$$

where

$$\mathbf{J}_B = \begin{bmatrix} \mathbf{B}_1 & \tilde{\mathbf{B}}_2 \\ \mathbf{I}_3 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} b_{11}^1 & b_{12}^1 & b_{13}^1 & b_{11}^2 & b_{12}^2 & 0 \\ b_{21}^1 & b_{22}^1 & b_{23}^1 & b_{21}^2 & b_{22}^2 & 0 \\ b_{31}^1 & b_{32}^1 & b_{33}^1 & b_{31}^2 & b_{32}^2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Simple matrix algebra gives

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix} \begin{bmatrix} b_{11}^1 & b_{12}^1 & b_{13}^1 & b_{11}^2 & b_{12}^2 & 0 \\ b_{21}^1 & b_{22}^1 & b_{23}^1 & b_{21}^2 & b_{22}^2 & 0 \end{bmatrix} - \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} [b_{31}^1\ b_{32}^1\ b_{33}^1\ b_{31}^2\ b_{32}^2\ 0] = \underset{2\times6}{\mathbf{0}}$$

and

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \delta_1 b_{11}^1 & \delta_1 b_{12}^1 & \delta_1 b_{13}^1 & \delta_1 b_{11}^2 & \delta_1 b_{12}^2 & \delta_1 b_{13}^2 \\ \delta_2 b_{21}^1 & \delta_2 b_{22}^1 & 0 & \delta_2 b_{21}^2 & \delta_2 b_{22}^2 & 0 \end{bmatrix} +$$
$$- \begin{bmatrix} \delta_1 b_{31}^1 & \delta_1 b_{32}^1 & \delta_1 b_{33}^1 & \delta_1 b_{31}^2 & \delta_1 b_{32}^2 & 0 \\ \delta_2 b_{31}^1 & \delta_2 b_{32}^1 & \delta_2 b_{33}^1 & \delta_2 b_{31}^2 & \delta_2 b_{32}^2 & 0 \end{bmatrix} = \underset{2\times6}{\mathbf{0}}$$

Finally, the two sets of restrictions are

$$\begin{cases} \delta_1 b^1_{11} + \delta_1 b^1_{31} = 1 \\ \delta_1 b^1_{12} + \delta_1 b^1_{32} = 0 \\ \delta_1 b^1_{13} + \delta_1 b^1_{33} = 0 \\ \delta_1 b^2_{11} + \delta_1 b^2_{31} = 0 \\ \delta_1 b^2_{12} + \delta_1 b^2_{32} = 0 \end{cases} \quad \text{and} \quad \begin{cases} \delta_2 b^1_{21} + \delta_2 b^1_{31} = 0 \\ \delta_2 b^1_{22} + \delta_2 b^1_{32} = 1 \\ \delta_2 b^1_{23} + \delta_2 b^1_{33} = 0 \\ \delta_2 b^2_{21} + \delta_2 b^2_{31} = 0 \\ \delta_2 b^2_{22} + \delta_2 b^2_{32} = 0 \end{cases}$$

The first (second) set links the parameters of the first (second) spread equation to the parameters of the equation for Δr_t.

A.2 Inexact present value model - $\tilde{\boldsymbol{\theta}}_t \neq 0$

If $\tilde{\boldsymbol{\theta}}_t \sim MDS(\mathbf{0}, \boldsymbol{\Sigma}_\theta)$, taking expectations conditional on the information set I_{t-1} of (10) gives:

$$E_{t-1}\check{\mathbf{S}}_t = \mathbf{M}_\delta E_{t-1}\check{\mathbf{S}}_{t+1} + \mathbf{D}_\delta E_{t-1}\Delta r_{t+1}$$

where $E_{t-1}\tilde{\boldsymbol{\theta}}_t = 0$. In the illustrative example above (with $k = 2$), the following correspondence between the variables of the theoretical equation and those of the statistical model (12):

$$E_{t-1}\check{\mathbf{S}}_t = \mathbf{R}_S \mathbf{J}_B \mathbf{Y}^*_t$$

$$E_{t-1}\check{\mathbf{S}}_{t+1} = \mathbf{R}_S E_{t-1}\mathbf{Y}^*_{t+1} = \mathbf{R}_S \mathbf{J}^2_B \mathbf{Y}^*_t$$

$$E_{t-1}\Delta r_{t+1} = \mathbf{R}_r E_{t-1}\mathbf{Y}^*_{t+1} = \mathbf{R}_r \mathbf{J}^2_B \mathbf{Y}^*_t$$

where

$$\mathbf{R}_S := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{R}_r := [0\ 0\ 1\ 0\ 0\ 0]$$

The substitution of the previous expressions in (10), leads to the following equation:

$$\mathbf{R}_S \mathbf{J}_B \mathbf{Y}^*_t = \mathbf{M}_\delta \mathbf{R}_S \mathbf{J}^2_B \mathbf{Y}^*_t + \mathbf{D}_\delta \mathbf{R}_r \mathbf{J}^2_B \mathbf{Y}^*_t$$

$$(\mathbf{R}_S \mathbf{J}_B - \mathbf{M}_\delta \mathbf{R}_S \mathbf{J}^2_B - \mathbf{D}_\delta \mathbf{R}_r \mathbf{J}^2_B)\mathbf{Y}^*_t = \mathbf{0}$$

since $\mathbf{Y}^*_t \neq \mathbf{0}$ a.s., the following equation must hold:

$$\mathbf{R}_S \mathbf{J}_B - \mathbf{M}_\delta \mathbf{R}_S \mathbf{J}^2_B - \mathbf{D}_\delta \mathbf{R}_r \mathbf{J}^2_B = \mathbf{0}$$

that is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{J}_B - \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{J}^2_B - \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} [0\ 0\ 1\ 0\ 0\ 0]\mathbf{J}^2_B = \underset{2\times 6}{\mathbf{0}}$$

where

$$\mathbf{J}_B = \begin{bmatrix} \mathbf{B}_1 & \tilde{\mathbf{B}}_2 \\ \mathbf{I}_3 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} b^1_{11} & b^1_{12} & b^1_{13} & b^2_{11} & b^2_{12} & 0 \\ b^1_{21} & b^1_{22} & b^1_{23} & b^2_{21} & b^2_{22} & 0 \\ b^1_{31} & b^1_{32} & b^1_{33} & b^2_{31} & b^2_{32} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

and

$$\mathbf{J}_B^2 = \begin{bmatrix} (b_{11}^1)^2 + b_{12}^1 b_{21}^1 + b_{13}^1 b_{31}^1 + b_{11}^2 & b_{11}^1 b_{12}^1 + b_{12}^1 b_{22}^1 + b_{13}^1 b_{32}^1 + b_{12}^2 & b_{11}^1 b_{13}^1 + b_{12}^1 b_{23}^1 + b_{13}^1 b_{33}^1 & b_{11}^1 b_{11}^2 + b_{12}^1 b_{21}^2 + b_{13}^1 b_{31}^2 & b_{11}^1 b_{12}^2 + b_{12}^1 b_{22}^2 + b_{13}^1 b_{32}^2 & 0 \\ b_{21}^1 b_{11}^1 + b_{22}^1 b_{21}^1 + b_{23}^1 b_{31}^1 + b_{21}^2 & b_{21}^1 b_{12}^1 + (b_{22}^1)^2 + b_{23}^1 b_{32}^1 + b_{22}^2 & b_{21}^1 b_{13}^1 + b_{22}^1 b_{23}^1 + b_{23}^1 b_{33}^1 & b_{21}^1 b_{11}^2 + b_{22}^1 b_{21}^2 + b_{23}^1 b_{31}^2 & b_{21}^1 b_{12}^2 + b_{22}^1 b_{22}^2 + b_{23}^1 b_{32}^2 & 0 \\ b_{31}^1 b_{11}^1 + b_{32}^1 b_{21}^1 + b_{33}^1 b_{31}^1 + b_{31}^2 & b_{31}^1 b_{12}^1 + b_{32}^1 b_{22}^1 + b_{33}^1 b_{32}^1 + b_{32}^2 & b_{31}^1 b_{13}^1 + b_{32}^1 b_{23}^1 + (b_{33}^1)^2 & b_{31}^1 b_{11}^2 + b_{32}^1 b_{21}^2 + b_{33}^1 b_{31}^2 & b_{31}^1 b_{12}^2 + b_{32}^1 b_{22}^2 + b_{33}^1 b_{33}^2 & 0 \\ b_{11}^1 & b_{12}^1 & b_{13}^1 & b_{11}^2 & b_{12}^2 & 0 \\ b_{21}^1 & b_{22}^1 & b_{23}^1 & b_{21}^2 & b_{22}^2 & 0 \\ b_{31}^1 & b_{32}^1 & b_{33}^1 & b_{31}^2 & b_{32}^2 & 0 \end{bmatrix}$$

After some algebra, the two sets of restrictions are the following:

$$\begin{cases} b_{11}^1 - \delta_1 b_{11}^2 - \delta_1 b_{31}^2 - \delta_1 b_{11}^1 b_{31}^1 - \delta_1 b_{12}^1 b_{21}^1 - \delta_1 b_{21}^1 b_{32}^1 - \delta_1 b_{13}^1 b_{31}^1 - \delta_1 b_{31}^1 b_{33}^1 - \delta_1 (b_{11}^1)^2 = 0 \\ b_{12}^1 - \delta_1 b_{12}^2 - \delta_1 b_{32}^2 - \delta_1 b_{11}^1 b_{12}^1 - \delta_1 b_{12}^1 b_{22}^1 - \delta_1 b_{12}^1 b_{31}^1 - \delta_1 b_{13}^1 b_{32}^1 - \delta_1 b_{22}^1 b_{32}^1 - \delta_1 b_{32}^1 b_{33}^1 = 0 \\ -\delta_1 b_{33^2}^1 - \delta_1 b_{13}^1 b_{33}^1 + b_{13}^1 - \delta_1 b_{11}^1 b_{13}^1 - \delta_1 b_{12}^1 b_{23}^1 - \delta_1 b_{13}^1 b_{31}^1 - \delta_1 b_{23}^1 b_{32}^1 = 0 \\ b_{11}^2 - \delta_1 b_{11}^1 b_{11}^2 - \delta_1 b_{12}^1 b_{21}^2 - \delta_1 b_{31}^1 b_{11}^2 - \delta_1 b_{13}^1 b_{31}^2 - \delta_1 b_{32}^1 b_{21}^2 - \delta_1 b_{33}^1 b_{31}^2 = 0 \\ b_{12}^2 - \delta_1 b_{11}^1 b_{12}^2 - \delta_1 b_{12}^1 b_{22}^2 - \delta_1 b_{31}^1 b_{12}^2 - \delta_1 b_{13}^1 b_{32}^2 - \delta_1 b_{32}^1 b_{22}^2 - \delta_1 b_{33}^1 b_{32}^2 = 0 \end{cases}$$

$$\begin{cases} b_{21}^1 - \delta_2 b_{21}^2 - \delta_2 b_{31}^2 - \delta_2 b_{11}^1 b_{21}^1 - \delta_2 b_{11}^1 b_{31}^1 - \delta_2 b_{21}^1 b_{22}^1 - \delta_2 b_{21}^1 b_{32}^1 - \delta_2 b_{31}^1 b_{23}^1 - \delta_2 b_{31}^1 b_{33}^1 = 0 \\ b_{22}^1 - \delta_2 b_{22}^2 - \delta_2 b_{32}^2 - \delta_2 b_{12}^1 b_{21}^1 - \delta_2 b_{12}^1 b_{31}^1 - \delta_2 b_{22}^1 b_{32}^1 - \delta_2 b_{23}^1 b_{32}^1 - \delta_2 b_{32}^1 b_{33}^1 - \delta_2 (b_{22}^1)^2 = 0 \\ -\delta_2 b_{33^2}^1 - \delta_2 b_{23}^1 b_{33}^1 + b_{23}^1 - \delta_2 b_{21}^1 b_{13}^1 - \delta_2 b_{13}^1 b_{31}^1 - \delta_2 b_{22}^1 b_{23}^1 - \delta_2 b_{23}^1 b_{32}^1 = 0 \\ b_{21}^2 - \delta_2 b_{21}^1 b_{11}^2 - \delta_2 b_{31}^1 b_{11}^2 - \delta_2 b_{22}^1 b_{21}^2 - \delta_2 b_{32}^1 b_{21}^2 - \delta_2 b_{23}^1 b_{31}^2 - \delta_2 b_{33}^1 b_{31}^2 = 0 \\ b_{22}^2 - \delta_2 b_{21}^1 b_{12}^2 - \delta_2 b_{31}^1 b_{12}^2 - \delta_2 b_{22}^1 b_{22}^2 - \delta_2 b_{32}^1 b_{22}^2 - \delta_2 b_{23}^1 b_{32}^2 - \delta_2 b_{33}^1 b_{32}^2 = 0 \end{cases}$$

In general, for a model with n $EHTS$ relations and k lags, the restrictions implied are $n(pk) - n$.
Moreover, considering an $MA(q)$ model for $\tilde{\boldsymbol{\theta}}_t$, the restrictions to be applied are given by the following equation, as derived by González and Gonzalo (2000):

$$\mathbf{R}^s \mathbf{J}_B^{q+1} - \mathbf{M}_\delta \mathbf{R}^s \mathbf{J}_B^{q+2} - \mathbf{D}_\delta \mathbf{R}^r \mathbf{J}_B^{q+2} = \mathbf{0}$$

B Results for other samples

B.1 Restricted sample

Table 6: **Test of Expectation Hypothesis - Summary results H_r, H_β, H_{CER}**

		Rank test - H_r			Beta test - H_β			CER test - H_{CER}				
								Exact model			Non exact model	
		trace stat.	p-value	bootstrap p-value	LR stat.	p-value	bootstrap p-value	Wald(HAC) stat	p-value	bootstrap p-value	Wald(HAC) stat	p-value
Bivariate System	*1W*	1.97	0.781	0.792	0.01	0.910	0.913	30.38	0.000	0.000	13.67	0.134
	2W	2.08	0.759	0.774	1.24	0.265	0.221	30.57	0.000	0.000	8.17	0.517
	3W	2.37	0.706	0.737	2.58	0.108	0.100	22.05	0.009	0.002	6.09	0.731
	1M	2.37	0.705	0.784	2.09	0.148	0.134	10.84	0.287	0.271	3.84	0.921
	2M	2.79	0.627	0.782	2.73	0.099	0.067	56.21	0.000	0.000	7.00	0.637
	3M	2.55	0.672	0.775	2.30	0.130	0.107	78.71	0.000	0.000	15.45	0.079
Joint System		2.33	0.886	0.676	9.432	0.151	0.196	11349.130	0.000			0.000

Notes: The table reports the results of the rank test (H_r), the test on the beta coefficients (H_β) and the test of the cross-equation restrictions (H_{CER}) (see Tables 1-3 for further details). Asymptotic and bootstrap p-values are reported.

B.2 Della Corte et al. (2008) sample

Table 7: **Test of Expectation Hypothesis - Summary results H_r, H_β, H_{CER}**

		Rank test - H_r			**Beta test - H_β**			**CER test - H_{CER}**				
								Exact model			**Non exact model**	
		trace stat.	**p-value**	**bootstrap p-value**	**LR stat.**	**p-value**	**bootstrap p-value**	**Wald(HAC) stat**	**p-value**	**bootstrap p-value**	**Wald(HAC) stat**	**p-value**
Bivariate System	*1W*	1.79	0.813	0.850	0.01	0.910	0.913	54.28	0.000	0.000	23.97	0.004
	2W	1.72	0.824	0.844	1.24	0.265	0.221	55.47	0.000	0.000	10.40	0.319
	3W	1.70	0.828	0.826	2.58	0.108	0.100	67.70	0.000	0.000	13.90	0.126
	1M	1.72	0.826	0.826	2.09	0.148	0.134	31.26	0.000	0.000	9.02	0.435
	2M	1.65	0.836	0.890	2.73	0.099	0.067	130.84	0.000	0.000	16.38	0.059
	3M	1.63	0.841	0.890	2.30	0.130	0.107	158.18	0.000	0.000	22.99	0.006
Joint System		1.93	0.926	0.820	12.995	0.043	0.040	9425.370	0.000			0.000

Notes: The table reports the results of the rank test (H_r), the test on the beta coefficients (H_β) and the test of the cross-equation restrictions (H_{CER}) (see Tables 1-3 for further details). Asymptotic and bootstrap p-values are reported.

References

Adrian, T. and Shin, H. S. (2010). Liquidity and leverage. *Journal of Financial Intermediation*, 19(3):418–437.

Bekaert, G. and Hodrick, R. J. (2001). Expectations hypotheses tests. *The Journal of Finance*, 56(4):1357–1394.

Boswijk, H. P., Cavaliere, G., Rahbek, A., and Taylor, A. (2013). Inference on co-integration parameters in heteroskedastic vector autoregressions. Technical report, Tinbergen Institute Discussion Paper.

Bårdsen, G. and Fanelli, L. (2014). Frequentist evaluation of small dsge models. *Journal of Business & Economic Statistics*, forthcoming.

Campa, J. M. and Gavilan, A. (2011). Current accounts in the euro area: An intertemporal approach. *Journal of International Money and Finance*, 30(1):205 – 228.

Campbell, J. Y. and Shiller, R. J. (1987). Cointegration and tests of present value models. *Journal of Political Economy*, 95(5):1062–88.

Cavaliere, G., Rahbek, A., and Taylor, A. (2014). Bootstrap determination of the co-integration rank in heteroskedastic var models. *Econometric Reviews*, (forthcoming).

Cavaliere, G., Rahbek, A., and Taylor, A. M. R. (2012). Bootstrap determination of the co-integration rank in vector autoregressive models. *Econometrica*, 80(4):1721–1740.

Della Corte, P., Sarno, L., and Thornton, D. L. (2008). The expectation hypothesis of the term structure of very short-term rates: Statistical tests and economic value. *Journal of Financial Economics*, 89(1):158–174.

Engel, C. and West, K. D. (2005). Exchange Rates and Fundamentals. *Journal of Political Economy*, 113(3):485–517.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

Fanelli, L. (2008). Testing the new keynesian phillips curve through vector autoregressive models: Results from the euro area*. *Oxford Bulletin of Economics and Statistics*, 70(1):53–66.

Fleming, M. J. and Garbade, K. D. (2002). When the back office moved to the front burner: settlement fails in the treasury market after 9/11. *Economic Policy Review*, (Nov):35–57.

González, M. and Gonzalo, J. (2000). Econometric implications of non-exact present value models.

Gorton, G. and Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, 104(3):425 – 451.

Hafner, C. M. and Herwartz, H. (2009). Testing for linear vector autoregressive dynamics under multivariate generalized autoregressive heteroskedasticity. *Statistica Neerlandica*, 63(3):294–323.

Hansen, P. R. (2003). Structural changes in the cointegrated vector autoregressive model. *Journal of Econometrics*, 114(2):261 – 295.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johansen, S. and Swensen, A. R. (2011). On a graphical technique for evaluating some models involving rational expectations. *Journal of Time Series Econometrics*, 3(1):1–29.

Krishnamurthy, A. (2009). How debt markets have malfunctioned in the crisis. Technical report, National Bureau of Economic Research.

Krishnamurthy, A., Nagel, S., and Orlov, D. (2012). Sizing up repo. Technical report, National Bureau of Economic Research.

Longstaff, F. A. (2000). The term structure of very short-term rates: New evidence for the expectations hypothesis. *Journal of Financial Economics*, 58(3):397–415.

Sarno, L., Thornton, D. L., and Valente, G. (2007). The empirical failure of the expectations hypothesis of the term structure of bond yields. *Journal of Financial and Quantitative Analysis*, 42(01):81–100.

Savin, N. (1984). Multiple hypothesis testing. In GrilichesĘ, Z. and Intriligator, M. D., editors, *Handbook of Econometrics*, volume 2 of *Handbook of Econometrics*, chapter 14, pages 827–879. Elsevier.

Sbordone, A. M. (2005). Do expected future marginal costs drive inflation dynamics? *Journal of Monetary Economics*, 52(6):1183–1197.

Shiller, R. J. (1979). The volatility of long-term interest rates and expectations models of the term structure. *The Journal of Political Economy*, pages 1190–1219.

Bootstrap rank determination in VAR models
A Stata command

Vanessa Gunnella[1]

Abstract

This paper introduces the stata command **bootrank** which implements the bootstrap likelihood ratio rank test algorithm developed by Cavaliere et al. (2012). The test improves the small sample properties of Johansen (1995) by generating $I(1)$ bootstrap samples under the null cointegration rank. Moreover, Cavaliere et al. (2014) show that the wild resampling scheme is both correctly sized and consistent under conditional and unconditional heteroskedasticity. The procedure to test $H(r)$ against $H(p)$ is implemented, as well as the sequential procedure. The test is consistent and it shows good size properties.

[1]University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Email: vanessa.gunnella2@unibo.it.
I wish to thank Giuseppe Cavaliere for useful comments and suggestions. The .ado file is available upon request.

1 Introduction

The rank test plays a crucial role in the cointegrated vector autoregressive (VAR) analysis. It allows to establish whether the endogenous variables of the model are cointegrated and, in case they are, it makes inference about the number of cointegrating relationships and the number of stochastic common trends driving the system.
Johansen (1988, 1991) treats the maximum likelihood estimation of reduced rank regression as an eigenvalues problem and in this framework he derives his likelihood ratio tests. The trace test is a likelihood ratio (LR)-type test. It compares the restricted model under the null $H(r)$, which imposes rank equal to r ($0 \leq r < p$), with the unrestricted model which assumes full rank, $H(p)$. In order to determine the cointegration rank, a sequential procedure can be implemented. Starting from $r = 0$, the test is carried out testing $H(r)$ against $H(p)$, until the non-rejection of the null hypothesis. The asymptotic distribution of the test depends on the number of common stochastic trends $(p-r)$ of the model and on the specification of the deterministic components. However, it has been shown (Johansen, 2002) that the asymptotic test suffers from poor small sample properties. For this reason bootstrap algorithms have been proposed in order to improve the test performance. These procedures are based either on *iid* resampling (Giersbergen, 1996; Swensen, 2006; Trenkler, 2009) or wild bootstrap (Cavaliere et al., 2010a,b). Swensen (2006) algorithm generate explosive bootstrap samples when the tested rank is smaller than the true rank and this compromises both finite and asymptotic properties of the test.
The bootstrap scheme proposed by Cavaliere et al. (2012) solves this issue using the estimated parameters and the residuals of the model under the null $H(r)$ for the generation of the bootstrap samples that are $I(1)$, as they are supposed to be. The authors prove that the estimates are asymptotically consistent even when $r < r_0$ and show through Monte Carlo experiments better size and power with respect to both the asymptotic test and Swensen (2006) procedure. Moreover, Cavaliere et al. (2014) also show that the *iid* bootstrap resampling scheme is correctly sized and consistent in the case of time-varying conditional variance. The alternative wild bootstrap they introduce, instead, preserves both properties also in the case on unconditional heteroskedasticity.

2 The bootstrap rank test

The bootstrap scheme assumes the following reduced rank VAR model for the multivariate process $\{\mathbf{X}_t\}_{t=1,..,\infty}$ of dimension p:

$$\mathbf{\Delta X}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{X}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \mathbf{\Delta X}_{t-i} + \boldsymbol{\alpha}\boldsymbol{\rho}'\mathbf{D}_t + \boldsymbol{\phi}\mathbf{d}_t + \boldsymbol{\varepsilon}_t \tag{1}$$

where $\mathbf{D}_t$ is the vector of deterministic components that lie in the cointegrating space and $\mathbf{d}_t$ is the vector of short run deterministics.
The bootstrap algorithm is based on the parameters estimated under the null hypothesis $H(r)$ so that it generates the bootstrap sample with the recursion:

$$\mathbf{\Delta X}^*_{r,t} = \hat{\boldsymbol{\alpha}}^{(r)}\hat{\boldsymbol{\beta}}^{(r)\prime}\mathbf{X}^*_{r,t-1} + \sum_{i=1}^{k-1} \hat{\mathbf{\Gamma}}_i^{(r)} \mathbf{\Delta X}^*_{t-i} + \hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\rho}}^{(r)\prime}\mathbf{D}_t + \hat{\boldsymbol{\phi}}^{(r)}\mathbf{d}_t + \boldsymbol{\varepsilon}^*_{r,t} \tag{2}$$

$\hat{\boldsymbol{\beta}}^{(r)}$ is a Gaussian QML estimator and $\hat{\boldsymbol{\alpha}}^{(r)}$, $\hat{\mathbf{\Gamma}}_i^{(r)}$ $(i = 1, .., k-1)$, $\hat{\boldsymbol{\rho}}^{(r)}$ and $\hat{\boldsymbol{\phi}}$ are obtained by OLS.
Hence the following two algorithms are proposed. The first one is to test $H(r)$ against $H(p)$, whereas the second one implements the sequential procedure.

ALGORITHM 1 :

1. Estimate equation 1 under $H(r)$ in order to get $\hat{\boldsymbol{\beta}}^{(r)}$, $\hat{\boldsymbol{\alpha}}^{(r)}$, $\hat{\mathbf{\Gamma}}_i^{(r)}$ $(i = 1, .., k-1)$, $\hat{\boldsymbol{\rho}}^{(r)}$ and $\hat{\boldsymbol{\phi}}^{(r)}$ and the residuals series $\hat{\boldsymbol{\varepsilon}}_{r,t}$.
2. Check the stability condition of the model, i.e. check whether the equation $\left|\hat{\mathbf{A}}^{(r)}(z)\right| = 0$ have $(p-r)$ roots on the unit circle and the remaining roots outside the unit circle.
3. Resample the centered residual $\hat{\boldsymbol{\varepsilon}}^c_{r,t}$ with:

   (a) *iid bootstrap*: generate an *iid* sequence of discrete uniform distribution with support $\{1, ..., T\}$, U_t, and associate the bootstrap residuals $\boldsymbol{\varepsilon}^*_{r,t}$ with $\hat{\boldsymbol{\varepsilon}}^c_{r,U_t}$

   (b) *wild bootstrap*: generate an *iid* $N(0, 1)$ sequence w_t and multiply it by the centered residual in order to obtain the bootstrap residuals $\boldsymbol{\varepsilon}^*_{r,t}$=$\hat{\boldsymbol{\varepsilon}}^c_{r,t}w_t$

   Then, generate $\mathbf{X}^*_t$ with the recursion 1, choosing as initial values those of the original series, i.e. $\mathbf{X}^*_{r,j} = \mathbf{X}_j,\ j = 1-k, ..., 0$

4. Compute the LR trace statistics, $Q^*_{r,T}$, as in Johansen (1995).

5. Repeat steps 1-4 B times in order to produce B conditionally independent bootstrap statistics. The p-value associated to the null $H(r)$ is obtained as $\tilde{p}^*_{r,T} := B^{-1}\sum_{b=1}^{B} \mathbb{1}(Q^*_{r,T:b} > Q_{r,T})$ where $Q_{r,T}$ is the trace statistics computed from the original sample. For $B \to \infty$ $\tilde{p}^*_{r,T} \to p^*_{r,T}$ with $p^*_{r,T} := 1 - G^*_{r,T}(Q_{r,T})$ and $G^*_{r,T}(Q_{r,T})$ conditional cdf of $Q^*_{r,T}$. The bootstrap test rejects $H(r)$ against $H(p)$ at level η if $p^*_{r,T} \leq \eta$.

ALGORITHM 2 : Starting from $r = 0$ perform the following steps:

1. - 4. Same as in ALGORITHM 1.

5. Same as in ALGORITHM 1 and if $p^*_{r,T} > \eta$ the estimated rank is r, otherwise repeat steps the proceduere testing the null $H(r+1)$ against $H(p)$ if $r+1 < p$ or the selected rank is p if $r+1 = p$.

3 The bootrank command

3.1 Syntax

bootrank *varlist* [*if*] [*in*], *lags*(#) *algorithm*(#) *bootstrap*(#) [trend(*string*) *ranksel*(#) *bootrep*(#)]

3.2 Options

lags (#) specifies the number to be included in the model, i.e. the order k of the VAR model in equation 1.

algorithm (#) selects the algorithm to be implemented. The first algorithm is run if 1 is typed, and the second is started if 2 is typed. If Algorithm 1 is selected, the optional option *ranksel* should be provided.

bootstrap (#) choices the resampling method for the bootstrap on the residuals: 1 for *iid* and 2 for wild gaussian bootstrap and 3 for wild Rademacher.

trend(none) specifies a model without trend or constant.

trend(constant) includes a constant in model.

trend(rconstant) includes an restricted constant in model.

trend(trend) includes a trend in model.

trend(rtrend) includes a restricted trend in model.

3.3 Optional options

ranksel (#) sets r, that is the rank under the null hypothesis for the first Algorithm.

bootrep (#) chooses the number of bootstrap replication B. If the option is not specified, the default number of replications is set to 499.

4 Empirical application

In this session Cavaliere et al. (2012) is illustrated through an empirical application on the term structure of interest rates in the US, as in **?**. The data are monthly time series of zero yields from the CRSP unsmoothed Fama and Bliss (1987) forward rates for monthly maturities $\tau = 3; 12; 36; 60; 120$. The sample goes from January 1970 to December 2009.
The dataset has to be **tsset**.

```
tsset time
      time variable:  time, 1970m1 to 2009m12
              delta:  1 month
```

Before proceeding with the rank test, the lag order k should be selected.

```
varsoc R3 R12 R36 R60 R120

 Selection-order criteria
 Sample:  1970m5 - 2009m12                    Number of obs      =       476
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-1588.41				.000556	6.69501	6.71222	6.73877
1	537.874	4252.6	25	0.000	8.1e-08	-2.13393	-2.0307	-1.8714*
2	611.381	147.01	25	0.000	6.6e-08	-2.33773	-2.14848*	-1.85644
3	638.698	54.634	25	0.001	6.6e-08*	-2.34747*	-2.07219	-1.6474
4	660.744	44.092*	25	0.011	6.7e-08	-2.33506	-1.97375	-1.41622

```
 Endogenous:  R3 R12 R36 R60 R120
  Exogenous:  _cons
```

Schwartz Bayesian Criteria suggests one lag, but the $VAR(1)$ displays some residual autocorrelation. Therefore, $k = 2$ is chosen, as suggested by Hannan-Queen information criterion.

4.1 Algorithm 1

First, the Algorithm 1 is chosen, in order to test $H(1)$ against $H(4)$.
For comparison purposes, the **bootrank** command performs Johansen (1995) asymptotic rank test and it reports the bootstrap p-values in the last column. Thus, after the $B = 499$ number of iterations performed, the command gives the following output:

```
. bootrank R3 R12 R36 R60 R120, lags(2) algorithm(1) bootstrap(1) trend(rconstant) ranksel(4) bootrep(499)

Bootstrapping rank test

r=4

   Eigenvalue stability condition

  +-------------------------------------+
  |        Eigenvalue       |  Modulus  |
  |-------------------------+-----------|
  |            1            |      1    |
  |   .9312203              |   .93122  |
  |   .9166932              |  .916693  |
  |   .8078905              |  .807891  |
  |   .6528032              |  .652803  |
  |  -.309308               |  .309308  |
  |   .2525407              |  .252541  |
  |  -.1480065              |  .148007  |
  |  -.1013044 + .07702823i |  .127263  |
  |  -.1013044 - .07702823i |  .127263  |
  +-------------------------------------+
   The VECM specification imposes a unit modulus.

                         Rank tests for cointegration
Trend: rconstant                                        Number of obs =     478
Sample:  1970m3 - 2009m12                                        Lags =       2
-------------------------------------------------------------------------------
                                                trace    Johansen    bootstrap
  rank    parms       LL       eigenvalue   statistic    5% cv        p-value
    4      53     613.58449     0.03044       2.6637      9.42       0.6754
```

First, the output of the **bootrank** command shows the eigenvalues, i.e. the reciprocal of the roots of the equation in point 2. of the Algorithm. If the condition of point 2. is violated, the error message "The VECM(r) is not stable" will appear.
The condition for the root of the characteristic polynomial are satisfied.
As it is possible to see, the bootstrap test cannot reject the null $H(4)$ at any significance level.

4.2 Algorithm 2

By selecting *algorithm*(2) and *bootstrap*(1) the sequential procedure with *iid* bootstrap resampling is performed. First, the conditions for the root of the characteristic polynomial are verified and reported for rank $r = 1, ..., 4$. Then, the sequential procedure reproduces the outcome of the test for $r = 1$ and it shows the test for the subsequent rank orders.
The outcome of the sequential test is the following:

```
. bootrank R3 R12 R36 R60 R120, lags(2) algorithm(2) bootstrap(1) trend(rconstant) bootrep(499)
```

Bootstrapping rank test

r=0

Eigenvalue stability condition

Eigenvalue	Modulus
-.4082793	.408279
-.208635 + .07717108i	.22245
-.208635 - .07717108i	.22245
.1779145	.177914
-.02019089	.020191

All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.

r=1

Eigenvalue stability condition

Eigenvalue	Modulus
1	1
1	1
1	1
1	1
.6431789	.643179
-.3735378	.373538
.2183904	.21839
-.154212 + .1007649i	.184214
-.154212 - .1007649i	.184214
-.05653751	.056538

The VECM specification imposes 4 unit moduli.

r=2

Eigenvalue stability condition

Eigenvalue	Modulus
1	1
1	1
1	1
.8043661	.804366
.6521013	.652101
-.319543	.319543
.2324492	.232449
-.1806897	.18069
-.09942323 + .08630866i	.131659
-.09942323 - .08630866i	.131659

The VECM specification imposes 3 unit moduli.

r=3

Eigenvalue stability condition

Eigenvalue	Modulus
1	1
1	1
.9081835	.908184
.8022318	.802232
.6537408	.653741
-.3076017	.307602
.2455766	.245577
-.1889533	.188953
-.08849081 + .08669478*i*	.123881
-.08849081 - .08669478*i*	.123881

The VECM specification imposes 2 unit moduli.

r=4

Eigenvalue stability condition

Eigenvalue	Modulus
1	1
.9312203	.93122
.9166932	.916693
.8078905	.807891
.6528032	.652803
-.309308	.309308
.2525407	.252541
-.1480065	.148007
-.1013044 + .07702823*i*	.127263
-.1013044 - .07702823*i*	.127263

The VECM specification imposes a unit modulus.

Rank tests for cointegration

Trend: rconstant — Number of obs = 478

Sample: 1970m3 - 2009m12 — Lags = 2

rank	parms	LL	eigenvalue	trace statistic	Johansen 5% cv	bootstrap p-value
0	25	519.09973		191.6332	76.07	0.0000
1	35	562.26629	0.16524	105.3001	53.12	0.0000
2	43	594.08745	0.12466	41.6577	34.91	0.0020
3	49	606.19699	0.04941	17.4387	19.96	0.1363
4	53	613.58449	0.03044	2.6637	9.42	0.6373
5	55	614.91632	0.00556			

The conditions for the root of the characteristic polynomial are verified for rank $r = 1, ..., 4$. As before, the p-value associated to $H(3)$ is not smaller than the 5% significance level, hence the null hypothesis of rank equal to 3 cannot be rejected.
Therefore, the sequential procedure with wild bootstrap $N(0, 1)$ is performed with the option *bootstrap*$(2)^2$:

```
. bootrank R3 R12 R36 R60 R120, lags(2) algorithm(2) bootstrap(2) trend(rconstant) bootrep(499)

Bootstrapping rank test
```

Rank tests for cointegration

Trend: rconstant — Number of obs = 478
Sample: 1970m3 - 2009m12 — Lags = 2

rank	parms	LL	eigenvalue	trace statistic	Johansen 5% cv	bootstrap p-value
0	25	519.09973		191.6332	76.07	0.0000
1	35	562.26629	0.16524	105.3001	53.12	0.0000
2	43	594.08745	0.12466	41.6577	34.91	0.0741
3	49	606.19699	0.04941	17.4387	19.96	0.3146
4	53	613.58449	0.03044	2.6637	9.42	0.8076
5	55	614.91632	0.00556			

As it is possible to see, the wild bootstrap supports the hypothesis of $r \geq 3$ only at 10% significance level.

[2] The stability check on the roots of the characteristic polynomial are the same as in the previous case.

References

Cavaliere, G., Rahbek, A., and Taylor, A. (2010a). Cointegration rank testing under conditional heteroskedasticity. *Econometric Theory*, 26(06):1719–1760.

Cavaliere, G., Rahbek, A., and Taylor, A. (2010b). Testing for co-integration in vector autoregressions with non-stationary volatility. *Journal of Econometrics*, 158(1):7–24.

Cavaliere, G., Rahbek, A., and Taylor, A. (2014). Bootstrap determination of the co-integration rank in heteroskedastic var models. *Econometric Reviews*, (forthcoming).

Cavaliere, G., Rahbek, A., and Taylor, A. M. R. (2012). Bootstrap determination of the co-integration rank in vector autoregressive models. *Econometrica*, 80(4):1721–1740.

Giersbergen, N. P. A. v. (1996). Bootstrapping the trace statistic in var models: Monte carlo results and applications. *Oxford Bulletin of Economics and Statistics*, 58(2):391–408.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2):231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–80.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johansen, S. (2002). A small sample correction for the test of cointegrating rank in the vector autoregressive model. *Econometrica*, 70(5):1929–1961.

Swensen, A. R. (2006). Bootstrap algorithms for testing and determining the cointegration rank in var models1. *Econometrica*, 74(6):1699–1714.

Trenkler, C. (2009). Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms. *Econometric Theory*, 25(01):243–269.