UvA ECONOMETRICS

Discussion Paper: 2013/08

Cointegration testing in panel VAR models under partial identification and spatial dependence

Artūras Juodis

www.ase.uva.nl/uva-econometrics

Amsterdam School of Economics
Department of Economics & Econometrics
Valckenierstraat 65-67
1018 XE AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Cointegration Testing in Panel VAR Models Under Partial Identification and Spatial Dependence☆

Artūras Juodis[a,b,*]

[a] *Amsterdam School of Economics, University of Amsterdam, Valckenierstraat 65-67, 1018 XE, Amsterdam, The Netherlands*
[b] *Tinbergen Institute, The Netherlands*

Abstract

This paper considers the Panel Vector Autoregressive Models of order 1 (PVAR(1)) with possibly spatially dependent error terms. We propose a simple Method of Moments based cointegration test using the rank test of Kleibergen and Paap (2006) for fixed number of time observations. The test is shown to be robust to spatial dependence, cross-sectional and time series heteroscedasticity as well as unbalanced panels. The main novelty of our approach is that we fully exploit the "weakness" of the Anderson and Hsiao (1982) moment conditions in construction of the new test. The finite-sample performance of the proposed test statistic is investigated using the simulated data. The results show that for most scenarios the method performs well in terms of both size and power. The proposed test is applied to employment and wage equations using Spanish firm data of Alonso-Borrego and Arellano (1999) and the results show little evidence for cointegration.

Keywords: Dynamic Panel Data, Panel VAR, cointegration, heteroscedasticity, spatial dependence, fixed T consistency.
JEL: C13, C33.

1. Introduction

The standard textbook treatment of econometrics assumes that estimation and hypothesis testing are the two sides of the same coin, with the latter being impossible to implement without the former. More importantly, for most of standard estimation methods regularity conditions necessary for hypothesis testing are equivalent to those of estimation. As the prototypical example consider the full rank assumption in the (Generalized) Method of Moments estimation that is needed both for estimation and hypothesis testing using Wald test. Unfortunately, for some econometric models this assumption can be too strong, resulting in a partial identification only, see e.g. Phillips (1989).

In their pioneering work Anderson and Rubin (1949) advocated the idea that it is possible to perform hypothesis testing for a simultaneous equations model under weaker regularity conditions than

☆I would like to thank Peter Boswijk and Maurice Bun for their comments and suggestions. I would also like to thank seminar participants at Tinbergen Institute, Netherlands Econometrics Study Group 2013 and "Conference on Cross-sectional Dependence in Panel Data" in Cambridge 2013. Financial support from the NWO MaGW grant "Likelihood-based inference in dynamic panel data models with endogenous covariates" is gratefully acknowledged.

*Tel. +31 626 51 8052
Email address: a.juodis@uva.nl (Artūras Juodis)

estimation[1]. Thus, in some situations it is possible to avoid the estimation step and to perform the hypothesis testing directly. In this paper we apply a similar principle by turning a disadvantageous situation from estimating point of view, into an advantageous one for hypothesis testing.

We consider the cointegration testing problem for the Panel VAR model of order 1 with fixed time dimension. To the best of our knowledge currently in the Dynamic Panel Data (DPD) literature with fixed number of time periods no feasible Method of Moments (MM) alternative to likelihood based cointegration testing procedures was proposed. The main reason for the absence of MM alternatives is the partial identification issue of the standard Anderson and Hsiao (1982)[AH] moment conditions when the process is cointegrated, as the Jacobian of these moment conditions are of reduced rank. Therefore we propose a rank based cointegration test for the Jacobian of the aforementioned moment conditions. We show that the proposed test is robust to cross-sectional and time series heteroscedasticity as well as spatial dependence between cross-sectional units. Moreover, unlike the likelihood based tests, (e.g. the Transformed Maximum Likelihood estimator of Binder, Hsiao, and Pesaran (2005)[henceforth BHP]) our test does not require any computationally demanding numerical optimization algorithms.

The paper is structured as follows. In Section 2 we briefly present the model and the intuition behind our testing procedure. The theoretical framework is presented in Section 3. In Section 4 we continue with the finite sample performance by means of a Monte Carlo analysis. In Section 5 we illustrate the testing procedure using the Spanish manufacturing data of Alonso-Borrego and Arellano (1999). Possible extensions and problems are discussed in Section 6. Finally, we conclude in Section 7.

2. Model

In this paper we consider the following PVAR(1) specification:

$$\boldsymbol{y}_{i,t} = \boldsymbol{\eta}_i + \boldsymbol{\Phi}\boldsymbol{y}_{i,t-1} + \boldsymbol{\varepsilon}_{i,t}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \tag{1}$$

where $\boldsymbol{y}_{i,t}$ is an $[m \times 1]$ vector, $\boldsymbol{\Phi}$ is an $[m \times m]$ matrix of parameters to be estimated, $\boldsymbol{\eta}_i$ is an $[m \times 1]$ vector of fixed effects and $\boldsymbol{\varepsilon}_{i,t}$ is an $[m \times 1]$ vector of innovations independent across i, with zero mean and covariance matrix $\boldsymbol{\Sigma}_{i,t}$. If we set $m = 1$ the model reduces to the linear DPD model with AR(1) dynamics. Throughout this paper we maintain the assumption that data in *levels* $\{\boldsymbol{y}_{i,t}\}$ is available for all $t = \{0, \dots, T\}$ and $i = \{1, \dots, N\}$.

We assume that $\boldsymbol{\eta}_i$ satisfy the so-called "common dynamics"("common factor") assumption:

$$\boldsymbol{\eta}_i = (\boldsymbol{I}_m - \boldsymbol{\Phi})\boldsymbol{\mu}_i.$$

If at least one eigenvalue is equal to unity this assumption ensures that there is no discontinuity in DGP, for further discussion see BHP.
Assuming common dynamics we can rewrite the model in (1) as:

$$\Delta\boldsymbol{y}_{i,t} = \boldsymbol{\Pi}\boldsymbol{u}_{i,t-1} + \boldsymbol{\varepsilon}_{i,t}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

[1]For more recent papers please refer Stock and Wright (2000) and Kleibergen (2005) *inter alia*.

Here we define $\boldsymbol{\Pi} = \boldsymbol{I}_m - \boldsymbol{\Phi}$ and $\boldsymbol{u}_{i,t-1} := \boldsymbol{y}_{i,t-1} - \boldsymbol{\mu}_i$. We say that series $\boldsymbol{y}_{i,t}$ are *cointegrated* if the $\boldsymbol{\Pi}$ matrix is of reduced rank[2]. In particular, there exist $[m \times r]$ matrices $\boldsymbol{\alpha}_r$ and $\boldsymbol{\beta}_r$[3] of full column rank such that:

$$\boldsymbol{\Phi} = \boldsymbol{I}_m + \boldsymbol{\alpha}_r \boldsymbol{\beta}_r',$$

where r is the rank of $\boldsymbol{\Pi}$. In general matrices $\boldsymbol{\alpha}_r$ and $\boldsymbol{\beta}_r$ are not unique as for any $[r \times r]$ invertible matrix $\boldsymbol{Q}$:

$$\boldsymbol{\alpha}_r \boldsymbol{\beta}_r' = \boldsymbol{\alpha}_r \boldsymbol{Q} \boldsymbol{Q}^{-1} \boldsymbol{\beta}_r' = \boldsymbol{\alpha}_r^* \boldsymbol{\beta}_r^{*\prime}.$$

This is the so-called *rotation problem*. As a result, it is a usual practise in the cointegration literature to impose identifying restrictions on $\boldsymbol{\alpha}_r$ or $\boldsymbol{\beta}_r$. However, the exact normalization of $\boldsymbol{\beta}_r$ matrix is not important for purpose of this paper as we are interested in rank of $\boldsymbol{\Pi}$ but not in particular form of $\boldsymbol{\Pi}$.

To explain the intuition of our approach lets consider the following (standard) AH moment conditions for Panel VAR(1) model:

$$\operatorname{vec} \mathrm{E}[(\Delta \boldsymbol{y}_{i,t} - \boldsymbol{\Phi} \Delta \boldsymbol{y}_{i,t-1}) \boldsymbol{y}_{i,t-2}'] = \boldsymbol{0}_{m^2}, \quad t = 2, \ldots, T.$$

The (minus) Jacobian of these moment conditions is given by:

$$\left(\mathrm{E}[\Delta \boldsymbol{y}_{i,t-1} \boldsymbol{y}_{i,t-2}'] \right)' \otimes \boldsymbol{I}_m, \quad t = 2, \ldots, T. \tag{2}$$

It follows from the properties of the Kronecker product that the rank of this matrix is determined by the rank of the matrix in the brackets[4]. The expected value of this term is given by (upon redefining $t \to t+1$, as the previous expression is well defined for $t = T+1$):

$$\mathrm{E}[\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}_{i,t-1}'] = \boldsymbol{\Pi} \, \mathrm{E}[\boldsymbol{u}_{i,t-1} \boldsymbol{y}_{i,t-1}'] + \mathrm{E}[\boldsymbol{\varepsilon}_{i,t} \boldsymbol{y}_{i,t-1}']. \tag{3}$$

Under usual regularity conditions of the DPD literature[5] the second term is equal to $\mathbf{O}_m$, while the first term is the product of rank r and rank m matrices. As a result $\operatorname{rk}(\mathrm{E}[\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}_{i,t-1}'])$ is equal to r and leads to violation of the "relevance" condition for the Instrumental Variable (IV) estimator. In such situation we can not estimate $\boldsymbol{\Phi}$ consistently from the AH moment conditions. However, we can use the Jacobian matrix directly avoiding the estimation step to test for cointegration.

3. Theoretical framework

3.1. Regularity conditions

In the previous section we have presented the intuition of the proposed method, but it still remains to be investigated under which conditions the $\mathrm{E}[\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}_{i,t-1}']$ term is of reduced rank iff $\boldsymbol{\Pi}$ is of reduced rank. We assume that the initial conditions $\boldsymbol{y}_{i,0}$ are of the following form:

$$\boldsymbol{y}_{i,0} = \boldsymbol{\Upsilon} \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{i,0}.$$

[2]Note that unlike pure time series models, we do not define cointegration as a property of time series.

[3]We slightly abuse the notation in this case, so that it remains consistent with the general practice of the time series cointegration literature.

[4]See e.g. Magnus and Neudecker (2007).

[5]That we will state formally in the next section.

Here $\boldsymbol{\Upsilon}$ is an $[m \times m]$ matrix that allows for a possible effect non-stationarity[6] of the initial condition if $\boldsymbol{\Upsilon} \neq \boldsymbol{I}_m$. The following *Standard Assumptions* are used throughout this paper[7].

(A.1) The error terms $\boldsymbol{\varepsilon}_{i,t}$ are i.h.d.[8] across cross sectional units and uncorrelated over time $\mathrm{E}[\boldsymbol{\varepsilon}_{i,t}\boldsymbol{\varepsilon}'_{i,s}] = \mathbf{O}_m$ for $s \neq t$. Variance of the error terms is a constant p.d. matrix $\operatorname{var} \boldsymbol{\varepsilon}_{i,t} = \boldsymbol{\Sigma}_{i,t}$ for $t > 0$. Furthermore, the higher order moment condition $\sup_{\leq i \leq N} \mathrm{E}[\|\boldsymbol{\varepsilon}_{i,t}\|^{4+\delta}] < \infty$ holds $\forall t$ and some $\delta > 0$.

(A.2) The fixed effects $\boldsymbol{\mu}_i$ are i.h.d. across cross sectional units and have zero mean with a p.d. variance-covariance matrix $\boldsymbol{\Sigma}_{i,\boldsymbol{\mu}}$. Furthermore, for all i and $t \geq 0$ $\mathrm{E}[\boldsymbol{\mu}_i\boldsymbol{\varepsilon}'_{i,t}] = \mathbf{O}_m$. The higher order moment condition $\sup_{\leq i \leq N} \mathrm{E}[\|\boldsymbol{\mu}_i\|^{4+\delta}] < \infty$ holds for some $\delta > 0$.

$\mathrm{E}[\boldsymbol{\varepsilon}_{i,t}\boldsymbol{y}'_{i,t-1}] = \mathbf{O}_m$ is a direct implication of Assumptions **(A.1)-(A.2)**. However, they do not ensure that $\boldsymbol{\Pi}\, \mathrm{E}[\boldsymbol{u}_{i,t-1}\boldsymbol{y}'_{i,t-1}]$ has a reduced rank iff $\boldsymbol{y}_{i,t}$ are cointegrated.
Lets investigate this issue more closely by expanding the $\mathrm{E}[\boldsymbol{u}_{i,t-1}\boldsymbol{y}'_{i,t-1}]$ term (for $t \geq 2$)[9]:

$$\begin{aligned}
\mathrm{E}[\boldsymbol{u}_{i,t-1}\boldsymbol{y}'_{i,t-1}] &= \mathrm{E}\left[\left(\boldsymbol{\Phi}^{t-1}\boldsymbol{u}_{i,0} + \sum_{s=0}^{t-2}\boldsymbol{\Phi}^s\boldsymbol{\varepsilon}_{i,t-s-1}\right)\left(\boldsymbol{\mu}_i + \boldsymbol{\Phi}^{t-1}\boldsymbol{u}_{i,0} + \sum_{s=0}^{t-2}\boldsymbol{\Phi}^s\boldsymbol{\varepsilon}_{i,t-s-1}\right)'\right] \\
&= \mathrm{E}\left[\left(\boldsymbol{\Phi}^{t-1}(\boldsymbol{\Upsilon} - \boldsymbol{I}_m)\boldsymbol{\mu}_i + \sum_{s=0}^{t-1}\boldsymbol{\Phi}^s\boldsymbol{\varepsilon}_{i,t-s-1}\right)\left((\boldsymbol{I}_m + \boldsymbol{\Phi}^{t-1}(\boldsymbol{\Upsilon} - \boldsymbol{I}_m))\boldsymbol{\mu}_i + \sum_{s=0}^{t-1}\boldsymbol{\Phi}^s\boldsymbol{\varepsilon}_{i,t-s-1}\right)'\right] \\
&= \underbrace{\boldsymbol{\Phi}^{t-1}(\boldsymbol{\Upsilon} - \boldsymbol{I}_m)\boldsymbol{\Sigma}_{\boldsymbol{\mu}}(\boldsymbol{\Phi}^{t-1}(\boldsymbol{\Upsilon} - \boldsymbol{I}_m))'}_{p.s.d} + \boldsymbol{\Phi}^{t-1}(\boldsymbol{\Upsilon} - \boldsymbol{I}_m)\boldsymbol{\Sigma}_{\boldsymbol{\mu}} \\
&\quad + \underbrace{\sum_{s=0}^{t-2}\boldsymbol{\Phi}^s\boldsymbol{\Sigma}_{t-1-s}\boldsymbol{\Phi}^{s'}}_{p.d.} + \mathrm{E}[\boldsymbol{\Phi}^{t-1}\boldsymbol{\varepsilon}_{i,0}\boldsymbol{\varepsilon}'_{i,0}(\boldsymbol{\Phi}^{t-1})']
\end{aligned}$$

In the effect-stationary case ($\boldsymbol{\Upsilon} = \boldsymbol{I}_m$) all terms involving $\boldsymbol{\Upsilon}$ are equal to $\mathbf{O}_m$. However, if that is not the case we have that the first term is a p.s.d. matrix while it is not immediately clear what happens with the second term. The third term is a p.d. matrix as all $\boldsymbol{\Sigma}_s$'s are positive definite. The analysis of the last term is more subtle as it requires explicit assumptions regarding the DGP for $\boldsymbol{\varepsilon}_{i,0}$. In general, we are looking at $\boldsymbol{\varepsilon}_{i,0}$ such that at least product matrix $\mathrm{E}[\boldsymbol{\Phi}^{t-1}\boldsymbol{\varepsilon}_{i,0}\boldsymbol{\varepsilon}'_{i,0}(\boldsymbol{\Phi}^{t-1})']$ is well defined and non-negative definite. Below we summarize few DGP's for $\boldsymbol{\varepsilon}_{i,0}$ currently used in the literature that satisfy this condition.

(DGP.1) $\boldsymbol{\varepsilon}_{i,0} \sim IID(\mathbf{0}, \boldsymbol{\Sigma}_0)$ with $\boldsymbol{\Sigma}_0$ constant p.s.d. matrix (independent of other DGP parameters).

(DGP.2) $\boldsymbol{\varepsilon}_{i,0} = \sum_{l=0}^{M}\boldsymbol{\Phi}^l\boldsymbol{\varepsilon}_{i,-l}$. Here M is assumed to be finite.

[6] It is often called "mean non-stationarity" by other scholars.
[7] If we allow for individual specific $\boldsymbol{\Upsilon}_i$ it is sufficient to assume that $\sup_{1\leq i\leq N}\|\boldsymbol{\Upsilon}_i\|^{4+\delta} < \infty$ to ensure that results hold without any changes.
[8] Here i.h.d. stands for "Independently and heterogeneously distributed".
[9] For the ease of exposition in what follows we assume that $\boldsymbol{\Sigma}_{i,t} = \boldsymbol{\Sigma}_t$ for all i and t.

(DGP.3) $\boldsymbol{\varepsilon}_{i,0} = \sum_{l=0}^{\infty} \left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right) \boldsymbol{\varepsilon}_{i,-l} + \boldsymbol{C}\boldsymbol{\xi}_i$. Here $\boldsymbol{\xi}_i$ is an $[m \times 1]$ vector of the (independent) individual-specific initialization effects, while $\boldsymbol{C} = \boldsymbol{\beta}_{\perp} \left(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp}\right)^{-1} \alpha'_{\perp}$ is an $m - r$ rank matrix[10].

To simplify matters we assume that all random variables in **(DGP.1)-(DGP.3)** satisfy assumptions **(A.1)-(A.2)**. The **(DGP.3)** initialization was used in the Monte Carlo studies of BHP and is motivated by the *Granger Representation Theorem*, see e.g. Johansen (1995)[Theorem 4.2]. The **(DGP.2)**, among others, was used in Hayakawa (2011).

It is important to emphasize that *all* three DGP's are well defined for all rank values of r. For $\rho(\boldsymbol{\Phi}) < 1$ we have $\boldsymbol{C} = \mathbf{O}_m$ resulting in stationary initialization. On the other hand, $\boldsymbol{\Phi} = \boldsymbol{I}_m$ implies $\boldsymbol{C} = \boldsymbol{I}_m$ (by definition) so that **(DGP.3)** and **(DGP.2)** coincide (by redefining M to $M + 1$).
For **(DGP.3)** by construction of the $\boldsymbol{C}$ matrix we have that $\boldsymbol{\Pi}\boldsymbol{C} = \mathbf{O}_m$, and thus $\boldsymbol{\Phi}^{t-1}\boldsymbol{C} = \boldsymbol{C}$, we have that:

$$\mathrm{E}[\boldsymbol{\Pi}\boldsymbol{\Phi}^{t-1}\boldsymbol{\varepsilon}_{i,0}\boldsymbol{\varepsilon}'_{i,0}(\boldsymbol{\Phi}^{t-1})'] = \boldsymbol{\Pi}\boldsymbol{\Phi}^{t-1}\left(\sum_{l=0}^{\infty}\left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right)\boldsymbol{\Sigma}\left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right)'\right)(\boldsymbol{\Phi}^{t-1})',$$

where existence of $\sum_{l=0}^{\infty}\left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right)\boldsymbol{\Sigma}\left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right)'$ is implied by the absolute summability of $\{\boldsymbol{\Phi}^l - \boldsymbol{C}\}_{l=0}^{\infty}$, see e.g. Lütkepohl (2006). Furthermore, it is obvious that $\sum_{l=0}^{\infty}\left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right)\boldsymbol{\Sigma}\left(\boldsymbol{\Phi}^l - \boldsymbol{C}\right)'$ is a p.s.d. matrix and consecutively that $\left(\mathrm{E}[\boldsymbol{\Phi}^{t-1}\boldsymbol{\varepsilon}_{i,0}\boldsymbol{\varepsilon}'_{i,0}(\boldsymbol{\Phi}^{t-1})']\right)$ is a p.s.d. matrix.

If we can ensure that $\boldsymbol{\Phi}^{t-1}(\boldsymbol{\Upsilon} - \boldsymbol{I}_m)\boldsymbol{\Sigma}_{\boldsymbol{\mu}}$ is such that $\mathrm{E}[\boldsymbol{u}_{i,t-1}\boldsymbol{y}'_{i,t-1}]$ has full rank m, then $\mathrm{E}[\boldsymbol{\Pi}\boldsymbol{u}_{i,t-1}\boldsymbol{y}'_{i,t-1}]$ has reduced rank r iff $\boldsymbol{y}_{i,t-1}$ are cointegrated[11]. However, it is not a trivial task to identify the parameter space of $\{\boldsymbol{\Phi}, \boldsymbol{\Upsilon}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}}\}$ for the aforementioned condition to be satisfied. One special case is obtained for $\boldsymbol{\Upsilon} = \boldsymbol{I}_m$ (effect stationarity) with other matrices being unrestricted (at least finite). Unfortunately, there are a lot of evidence in the DPD literature suggesting that in general this assumption can be too restrictive, see e.g. Arellano (2003b) and Roodman (2009). In the Monte Carlo simulations we will check the adequacy of the proposed procedure by considering different values of $\boldsymbol{\Upsilon}$ that are mentioned in the literature.

3.2. Rank Test

In this paper we use the generalized rank test of Kleibergen and Paap (2006)[KP] as a basis for a cointegration testing. We will briefly introduce their testing procedure and later apply it to our problem. In construction of the rank test KP use the property that any $[k \times m]$ matrix $\boldsymbol{D}$ can be decomposed as:

$$\boldsymbol{D} = \boldsymbol{A}_q\boldsymbol{B}_q + \boldsymbol{A}_{q,\perp}\boldsymbol{\Lambda}_q\boldsymbol{B}_{q,\perp},$$

where all $\perp$ matrices are defined in a usual way and $\boldsymbol{\Lambda}_q$ is an $[(k - q) \times (m - q)]$ matrix. For $\boldsymbol{\Lambda}_q = \mathbf{O}$ the rank of $\boldsymbol{D}$ is determined by the rank of $\boldsymbol{A}_q\boldsymbol{B}_q$. The procedure in KP is based on testing if $\boldsymbol{\Lambda}_q$ is equal to $\mathbf{O}_m$, with matrices $\boldsymbol{A}_q, \boldsymbol{B}_q, \boldsymbol{\Lambda}_q$ obtained using the singular value decomposition (SVD). In our case matrix $\boldsymbol{D}$ is the $[m \times m]$ matrix $\mathrm{E}[\Delta\boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}]$.

[10] For simplicity in what follows, we assume that all $\boldsymbol{\varepsilon}_{i,-l}$ are homoscedastic over time.

[11] Note that positive definiteness of $\mathrm{E}[\boldsymbol{u}_{i,t-1}\boldsymbol{y}'_{i,t-1}]$ is a sufficient, but not a necessary condition. The term can be negative definite or even indefinite, as long as it has full rank.

We define the following cross-sectional average:

$$\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}} := \frac{1}{N}\sum_{i=1}^{N} \Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}.$$

Using the standard Lindeberg-Lévy CLT for i.h.d. data it follows that:

$$\sqrt{N}\,\mathrm{vec}\left(\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}} - \mathrm{E}[\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}]\right) \xrightarrow{d}, \mathbf{N}_{m^2}(\mathbf{0}_{m^2}, \boldsymbol{V}), \quad t = 2, \ldots, T.$$

Here the full rank matrix $\boldsymbol{V}$ can be consistently estimated using its finite sample counterpart:

$$\boldsymbol{V}_N = \frac{1}{N}\sum_{i=1}^{N} \mathrm{vec}\,(\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1})\,\mathrm{vec}\,(\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1})' - \mathrm{vec}\,\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}\,\mathrm{vec}\,\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}'.$$

Consecutively the estimator $\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}$ satisfies **Assumption 1** (asymptotic normal distribution), while $\boldsymbol{V}$ satisfies **Assumption 2**(full rank) of KP. These two assumptions are sufficient to Theorem **1** of KP to the problem at hand:

Theorem 1. *Let Assumptions **(A.1)-(A.2)** be satisfied with $\boldsymbol{\varepsilon}_{i,0}$ generated by one of **(DGP.1)-(DGP.3)**, then:*

$$\sqrt{N}\hat{\boldsymbol{\lambda}}_r \xrightarrow{d} \mathbf{N}(\mathbf{0}_r, \boldsymbol{\Omega}_r),$$

where:

$$\hat{\boldsymbol{\lambda}}_r = \mathrm{vec}\,\hat{\boldsymbol{\Lambda}}_r, \quad \hat{\boldsymbol{\Lambda}}_r = \hat{\boldsymbol{A}}'_{r,\perp}\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}\hat{\boldsymbol{B}}'_{r,\perp},$$
$$\boldsymbol{\Omega}_r = (\boldsymbol{B}_{r,\perp} \otimes \boldsymbol{A}'_{r,\perp})\boldsymbol{V}(\boldsymbol{B}_{r,\perp} \otimes \boldsymbol{A}'_{r,\perp})'$$

Furthermore, under $\mathbf{H} : \mathrm{rk}\,\boldsymbol{\Pi} = r$, the test statistic:

$$rk(r) = N\hat{\boldsymbol{\lambda}}'_r\boldsymbol{\Omega}_r^{-1}\hat{\boldsymbol{\lambda}}'_r$$

converges in distribution to a $\chi^2((m-r)^2)$ distributed random variable.

All $\boldsymbol{A}$ and $\boldsymbol{B}$ variables in Theorem 1 are obtained from the SVD of the $\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}$ matrix. An operational version of the $rk(r)$ test statistic is obtained by replacing the (unknown) matrix $\boldsymbol{\Omega}_r$ with some consistent estimator. An obvious choice for $\hat{\boldsymbol{\Omega}}_r$ is given by:

$$\hat{\boldsymbol{\Omega}}_r = (\hat{\boldsymbol{B}}_{r,\perp} \otimes \hat{\boldsymbol{A}}'_{r,\perp})\boldsymbol{V}_N(\hat{\boldsymbol{B}}_{r,\perp} \otimes \hat{\boldsymbol{A}}'_{r,\perp})'.$$

The test statistic in Theorem 1 is based only on one time series observation (in a sense that if $T > 2$, then we can construct test statistic for every value of t, but $t = 1$). Of course, it is not the most efficient way how information can be used. Instead, all time series observations can be pooled into one test statistic for testing rank of:

$$\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}_T = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T-1}\sum_{t=2}^{T} \Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}. \tag{4}$$

For any fixed value of T, the $\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}_T$ term satisfies sufficient conditions for the CLT , so that the results of Theorem 1 can be extended trivially, with $\boldsymbol{V}_N$ for this case given by:

$$\boldsymbol{V}_N = \frac{1}{N}\sum_{i=1}^{N} \mathrm{vec}\left(\frac{1}{T-1}\sum_{t=2}^{T}\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}\right)\mathrm{vec}\left(\frac{1}{T-1}\sum_{t=2}^{T}\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}\right)' - \mathrm{vec}\,\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}_T\,\mathrm{vec}\,\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}_T{}'.$$

In the next section we will use "rk-J" to denote the Jacobian based cointegration test for $\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}_T$.

Up to this stage we considered only the Jacobian of Anderson and Hsiao (1982) moment conditions, however for $T > 2$ further lags can be used. The particular choice of lags used is subject to the same "arbitrariness" as the choice of moment conditions for the Arellano and Bond (1991)[AB] estimator. More importantly, it is not clear that use of lags larger than 1 still ensures that $\mathrm{E}[\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-j}}_T]$ has reduced rank r iff $\mathrm{rk}\,\boldsymbol{\Pi} = r$ (even in the effect stationary case). Moreover, the power of the test might be substantially affected by the choice of lags, as with any alternative close to the unit circle we encounter weak instruments problem for any distanced lags. On the other hand, we can expect better test power to alternatives with substantially lower $\rho(\boldsymbol{\Phi})$.

Remark 1. If the model contains time effects λ_t, the test statistic is based on variables in deviations from the cross-sectional averages $\check{\boldsymbol{y}}_{i,t} := \boldsymbol{y}_{i,t} - (1/N)\sum_{i=1}^{N}\boldsymbol{y}_{i,t}$ rather than levels (similarly to the standard GMM treatment).

Remark 2. One important advantage of the proposed test statistic is the additional flexibility while dealing with unbalanced panels. As long as for every individual i at least one $\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}$ $(t > 1)$ term is available, the test statistic can be computed. The only difference as compared to the unbalanced case is that individual contributions to $\overline{\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}}_T$ are no longer simple averages with $T-1$ terms involved, but have individual specific number of observations $T(i) - 1$.

3.3. Cross-sectional dependence

In many situations of practical relevance the cross-sectional independence assumption might be too restrictive to properly describe the data at hand. In this section we generalize the testing procedure by allowing *weak* (for a general reference, see e.g. Sarafidis and Wansbeek (2012)) cross-sectional dependence in $\boldsymbol{\varepsilon}_{i,t}$. In particular, for simplicity we assume that $\boldsymbol{\varepsilon}_{i,t}$ have Spatial Moving Average (SMA) representation of the following form:

$$\boldsymbol{\varepsilon}_{i,t} = \theta\sum_{j=1}^{N}\omega_{i,j}\boldsymbol{\zeta}_{j,t} + \boldsymbol{\zeta}_{i,t}, \quad i = 1,\ldots,N; \quad t = 0,\ldots,T;$$

where all $\boldsymbol{\zeta}_{i,t}$ are independent across individuals and satisfy Assumptions **(A.1)-(A.2)**. Denote by $\boldsymbol{W}_N$ an $[N \times N]$ matrix with a typical element $\omega_{i,j}$. The spatial covariance structure induced by the SMA assumption is *local* (see Anselin (2003)), as compared to the *global* covariance structure induced by the Spatial Autoregressive model (SAR). Although we restrict our attention to the SMA model, the results can be similarly extended to both SAR and Spatial Error Components (SEC) specifications[12]. We assume that $\boldsymbol{W}_N$ and θ satisfy the following regularity conditions (see e.g. Kelejian and Prucha (2010) and Sarafidis (2011)):

[12] Assuming that regularity conditions for SAR and SEC models are satisfied as in Baltagi et al. (2007) or Kelejian and Prucha (2010).

(A.3) weighting matrix and space of MA parameter (i) All diagonal elements of $\boldsymbol{W}_N$ are equal to zero. (ii) The spatial moving average parameter satisfies $\theta \in (-c_{1,\theta}, c_{2,\theta})$ with $0 < c_{1,\theta}, c_{2,\theta} \leq c_\theta < \infty$. (iii) The matrix $\boldsymbol{W}_N$ is non-singular and $\boldsymbol{P}_N = \boldsymbol{I}_N + \theta \boldsymbol{W}_N$ is non-singular for all $\theta \in (-c_{1,\theta}, c_{2,\theta})$. (iv) The row and column sums of $\boldsymbol{W}_N$ and $\boldsymbol{P}_N$ are bounded uniformly in absolute value.

It can be easily seen that under Assumptions **(A.1)-(A.3)** the matrix $\mathrm{E}[\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}'_{i,t-1}]$ $(\forall i, t)$ has reduced rank r iff $\boldsymbol{\Pi}$ has reduced rank r (of course the same problem with effect non-stationary initial condition remains). Hence, the intuition does not change if spatial dependence across cross-sectional units is present. However, we can no longer use standard LLN and CLT for i.h.d. data to obtain limiting distribution of $\mathrm{vec}\,(\overline{\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}_{i,t-1}}_T)$. In order to prove the next result we strengthen Assumptions **(A.1)-(A.3)** by replacing all *uncorrelated* terms with *totally independent*, we denote these assumptions by **(A.1)*-(A.3)***.

Theorem 2. *Let Assumptions* ***(A.1)*-(A.3)**** *be satisfied and $\boldsymbol{\varepsilon}_{i,0}$ be generated by (one of)* ***(DGP.1)-(DGP.3)****, then:*

$$\frac{1}{\sqrt{N}} \mathrm{vec} \left(\sum_{i=1}^{N} \Delta \boldsymbol{y}_{i,t} \boldsymbol{y}'_{i,t-1} - \mathrm{E}[\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}'_{i,t-1}] \right) \xrightarrow{d}, \mathbf{N}_{m^2}(\mathbf{0}_{m^2}, \boldsymbol{V}), \quad t = 2, \ldots, T.$$

Proof. Observe that all elements in $\mathrm{vec}\,(\cdot)$ are at most quadratic in $\{\boldsymbol{\zeta}_{i,t}\}_{i=1,t=0}^{N,T}$ and $\{\boldsymbol{\mu}_i\}_{i=1}^N$ (call this stacked $[mN(T+2) \times 1]$ vector $\boldsymbol{\xi}$). All elements of $\boldsymbol{\xi}$ have zero mean with block-diagonal covariance matrix between individuals and satisfy Assumption **(A.2.)** of Kelejian and Prucha (2010). Furthermore, the quadratic form matrices $\{\boldsymbol{A}_{s,mN(T+2)}\}_{s=1}^{m^2}$ are functions of $(\boldsymbol{P}_N, \boldsymbol{\Phi}_0, \boldsymbol{\Upsilon})$ and consecutively have bounded column and row sums. The result follows from Theorem **A.1** in Kelejian and Prucha (2010). ☐

Remark 3. The implications of Theorem 2 are actually wider than considered in this paper. For instance, it enables use of the weak-instrument robust testing procedures of Kleibergen (2005) and Kleibergen and Mavroeidis (2009) for the AB estimator in PVAR(1) model with the SMA cross-sectional dependence.

So far we have considered models with *weak* cross-sectional dependence. However, there is a growing literature on DPD models with *strong* cross-sectional dependence (both for large and fixed T). The most popular model is the common factor type of dependence, see e.g. Pesaran and Tosseti (2011), Bai and Ng (2004) and recent surveys of Breitung and Pesaran (2008) and Sarafidis and Wansbeek (2012). The DGP for $\boldsymbol{y}_{i,t}$ can be expressed as following:

$$\boldsymbol{y}_{i,t} = \boldsymbol{\Phi} \boldsymbol{y}_{i,t-1} + \boldsymbol{\Lambda}_i \boldsymbol{f}_t + \boldsymbol{\varepsilon}_{i,t}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{5}$$

where $\boldsymbol{\Lambda}_i$ is an $[m \times k_f]$ matrix of *factor loadings* and $\boldsymbol{f}_t$ is an $[k_f \times 1]$ vector of common factors. The error terms $\boldsymbol{\varepsilon}_{i,t}$ can be i.h.d. or spatially dependent (both cases will be discussed separately). As the time dimension is assumed to be fixed we treat all $\boldsymbol{f}_t$ as fixed parameters, while the factor loadings $\boldsymbol{\Lambda}_i$ are assumed to satisfy:

$$\mathrm{vec}\, \boldsymbol{\Lambda}_i \sim (\mathbf{0}_{mk_f}, \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}).$$

For $\boldsymbol{f}_t = 1$ we have the fixed effects model. Subtracting $\boldsymbol{y}_{i,t-1}$ from both sides of (5):

$$\Delta \boldsymbol{y}_{i,t} = \boldsymbol{\Pi} \boldsymbol{y}_{i,t-1} + \boldsymbol{\Lambda}_i \boldsymbol{f}_t + \boldsymbol{\varepsilon}_{i,t}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{6}$$

The corresponding Jacobian of the AB moment conditions is then given by:

$$\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}'_{i,t-1} = \boldsymbol{\Pi} \boldsymbol{y}_{i,t-1} \boldsymbol{y}'_{i,t-1} + \boldsymbol{\Lambda}_i \boldsymbol{f}_t \boldsymbol{y}'_{i,t-1} + \boldsymbol{\varepsilon}_{i,t} \boldsymbol{y}'_{i,t-1}.$$

The second term on the RHS in general will not be equal to $\mathbf{O}_m$, hence the rk-J test can not be used without any changes. However, if the common dynamics is imposed on $\boldsymbol{\Lambda}_i$, then the previous equation can be rewritten as (upon redefinition $\boldsymbol{\Lambda}_i \to -\boldsymbol{\Pi} \boldsymbol{\Lambda}_i$):

$$\Delta \boldsymbol{y}_{i,t} \boldsymbol{y}'_{i,t-1} = \boldsymbol{\Pi} (\boldsymbol{y}_{i,t-1} - \boldsymbol{\Lambda}_i \boldsymbol{f}_t) \boldsymbol{y}'_{i,t-1} + \boldsymbol{\varepsilon}_{i,t} \boldsymbol{y}'_{i,t-1}.$$

If this Assumption is indeed satisfied, then the rk-J statistic does not require any adjustments. Unfortunately, there are several major drawbacks of imposing common dynamics in this case. First of all, unlike in the *pure* fixed effects case, motivation for this assumption is not clear. Secondly, even in the simple fixed effects case we have shown that it is difficult to ensure that $\mathrm{E}[\boldsymbol{u}_{i,t} \boldsymbol{y}'_{i,t-1}]$ is of full rank, this task becomes even more complicated under presence of common factor structure.
However, if $\boldsymbol{\varepsilon}_{i,t}$ are spatially correlated with some *known* weighting matrix $\boldsymbol{W}_N$ then we can make use of the so-called "spatial instruments" of Sarafidis (2011):

$$\hat{\boldsymbol{y}}'_{t-1} = (\boldsymbol{W}_N + \boldsymbol{W}'_N) \boldsymbol{y}'_{t-1}. \tag{7}$$

Here $\boldsymbol{y}'_{t-1}$ is an $[N \times m]$ matrix of stacked $\boldsymbol{y}'_{i,t-1}$ observations. The Jacobian of these moment conditions is then given by:

$$\Delta \boldsymbol{y}_{i,t} \hat{\boldsymbol{y}}'_{i,t-1} = \boldsymbol{\Pi} \boldsymbol{y}_{i,t-1} \hat{\boldsymbol{y}}'_{i,t-1} + \boldsymbol{\Lambda}_i \boldsymbol{f}_t \hat{\boldsymbol{y}}'_{i,t-1} + \boldsymbol{\varepsilon}_{i,t} \hat{\boldsymbol{y}}'_{i,t-1}.$$

As $\boldsymbol{\Lambda}_i$ are i.i.d. across cross-sectional units $\mathrm{E}[\boldsymbol{\Lambda}_i \boldsymbol{f}_t \hat{\boldsymbol{y}}'_{i,t-1}] = \mathrm{E}[\boldsymbol{\varepsilon}_{i,t} \hat{\boldsymbol{y}}'_{i,t-1}] = \mathbf{O}$, while the first term will be of reduced rank r iff $\mathrm{rk}\, \boldsymbol{\Pi} = r$ (for correctly specified $\boldsymbol{W}_N$ and $\theta \neq 0$). Of course, for small values of θ $\mathrm{E}[\boldsymbol{y}_{i,t-1} \hat{\boldsymbol{y}}'_{i,t-1}]$ will be very close to zero matrix resulting in the weak instrument problem, regardless of $\mathrm{rk}\, \boldsymbol{\Pi}$. Note that these moment conditions have merits even for situations without strong cross-sectional dependence as it directly solves the effect non-stationarity issue encountered in Section 3.
To sum up, we can see that for models with strong cross-sectional dependence the rk-J test can be useful only under strong assumptions about factor loadings and/or spatial dependence. In all other situations, we suspect that we should instead look for an estimation method that is robust to both cointegration and strong cross-sectional dependence so that the *rk* test of KP can be applied directly on $\hat{\boldsymbol{\Pi}}$ (for instance the estimator of Robertson et al. (2010)).

4. Monte Carlo

To the best of our knowledge only the BHP study provides results on cointegration analysis for panels with fixed T[13]. Hence, for the main building blocks of the finite-sample studies performed in

[13]Other studies, like Mutl (2009) adapted setups of BHP.

this paper we take the setups from BHP, but we provide extended range of scenarios. Only *bivariate* panels are considered, thus the only null hypothesis we are testing is:

$$\mathbf{H}_0 : \operatorname{rk} \boldsymbol{\Pi} = 1 \tag{8}$$

For simplicity we will use **(DGP.2)** for initialization:

$$\boldsymbol{y}_{i,0} = \boldsymbol{\Upsilon}_i \boldsymbol{\mu}_i + \sum_{j=0}^{M} \boldsymbol{\Phi}^j \boldsymbol{\varepsilon}_{i,-j}, \quad \boldsymbol{\varepsilon}_{i,-j} \sim IID(\mathbf{0}_2, \boldsymbol{\Sigma}). \tag{9}$$

In what follows we will alow for cross-sectional heterogeneity in $\boldsymbol{\Upsilon}_i$ but not in $\boldsymbol{\Sigma}$. We set $M = 50$ and number of Monte Carlo replications $B = 10000$[14].

We generate the individual heterogeneity $\boldsymbol{\mu}_i$ using exactly the same procedure as in BHP:

$$\boldsymbol{\mu}_i = \tau \left(\frac{q_i - 1}{\sqrt{2}} \right) \check{\boldsymbol{\eta}}_i, \quad q_i \sim \chi^2(1), \quad \check{\boldsymbol{\eta}}_i \sim \mathrm{N}(\mathbf{0}_2, \boldsymbol{\Sigma}). \tag{10}$$

We assume that the error terms are normally distributed i.i.d. both across individuals and time with zero mean and variance-covariance matrix $\boldsymbol{\Sigma}$ (to be specified later).

Before summarizing Design parameters for this Monte Carlo study recall that $\boldsymbol{\Pi}$ can rewritten as (for $m = 2$):

$$\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}' + \lambda \boldsymbol{\alpha}_\perp \boldsymbol{\beta}'_\perp$$

We set $\lambda = 0$ to study the size of the test, while non-zero values of λ are used to investigate power. In order to reduce the dimensionality of the parameter space we assume that vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are of the following structure:

$$\boldsymbol{\alpha} = \alpha \boldsymbol{\imath}_2, \quad \boldsymbol{\beta}' = (1, -0.2).$$

All Design parameters are summarized in Table 1[15].

Table 1: Design parameters. $d_i \sim Bernoulli(.3)$

N	T	τ	α	θ	λ	vech $\boldsymbol{\Sigma}$	$\boldsymbol{\Upsilon}$
50	3	1	-.1	.0	-.700	$(.05, .03, .05)'$	$0.5\boldsymbol{I}_2$
250	5	5	-.5	.3	-.300	$(.05, -.03, .05)'$	$\boldsymbol{I}_2$
500	7		-.9	.5	-.100		$1.5\boldsymbol{I}_2$
					-.050		$\boldsymbol{I}_2 - \boldsymbol{\Phi}^{10}$
					-.010		$\boldsymbol{I}_2 - d_i\boldsymbol{\Phi}^{10}$
					-.005		$\begin{pmatrix} .85 & .15 \\ .00 & .85 \end{pmatrix}$
					.000		

Comparing our Designs to those present in the literature, we can see that Design **3** of BHP is achieved when $\alpha = -0.5$.

[14]As it was argued in Kiviet (2012) 10000 Monte Carlo replications should be sufficient to get precise quantile estimates.

[15]Later we will use notation $\boldsymbol{\Upsilon}^{(i)}$ with i indicating the particular row of Table 1.

The θ parameter controls the degree of cross-sectional dependence between units. For $\theta = 0$ we have i.i.d. dataset, while for $\theta \neq 0$ cross-sectional units are *weakly* correlated. Spatial correlation matrix $\boldsymbol{W}_N$ is assumed to be **1 ahead - 1 behind** circular, so that every individual i is directly linked only with individuals $i-1$ and $i+1$.[16] Particular choice of the spatial matrix $\boldsymbol{W}_N$ is motivated by the study in Baltagi et al. (2007), where in the context of the panel unit root testing it is shown that the tests are mostly distorted for this choice of spatial matrix[17]. Thus, we suspect that by choosing this particular matrix we will put the proposed cointegration test under the least favorable conditions in terms of size distortions.

As we have discussed in Section 2 in the effect non-stationary case particular choice of $\{\boldsymbol{\Upsilon}, \boldsymbol{\Sigma}\}$ and τ might substantially influence the performance of the test statistic. For this reason we consider two different choices of $\boldsymbol{\Sigma}$ matrix.

The choice of $\boldsymbol{\Upsilon}^{(4)}$ is motivated by the finite start-up assumption, so that the individual specific effects are accumulated only for 9 periods. The particular choice of $S = 10$ was rather arbitrary and is not empirically or theoretically motivated[18]. With $\boldsymbol{\Upsilon}_i^{(5)} = \boldsymbol{I}_2 - d_i \boldsymbol{\Phi}^{10}$ we allow for mixture type of cross-sectional heterogeneity in $\boldsymbol{\Upsilon}_i$ such that approximately 30% of population has effect non-stationary initial condition, while the rest have effect stationary one. $\boldsymbol{\Upsilon}^{(6)}$ is based on the estimates in Arellano (2003a) obtained from the bivariate panel of Spanish firm data.

In terms of the test power, we suspect that it should be decreasing with $|\lambda|$, with almost no power against alternatives with $\lambda \approx 0$. However, it is very likely that for general $\boldsymbol{\Upsilon}$ matrices power curve might not be monotonic because λ not only controls the rank of $\boldsymbol{\Pi}$ but as well (indirectly) the eigenvalues of the $\mathrm{E}[\boldsymbol{u}_{i,t-1}\boldsymbol{y}_{i,t-1}']$ matrix. Hence, for some specific choices of $\boldsymbol{\Upsilon}$ we can observe the weak instruments problem of the AB moment conditions that is not caused by the reduced rank of $\boldsymbol{\Pi}$ matrix.

Remark 4. It is important to emphasize that in this MC study we *do not* control $\mathbf{R}^2_\Delta$(or any other sensible measure as in Kiviet (2007)) between designs as it was done in BHP. It is not clear what exactly should be kept fixed between designs as for the effect non-stationary designs measures will be time dependent (the average of $\mathbf{R}^2_\Delta$ over time is a likely candidate). We acknowledge that by not fixing relevant "orthogonal" measures we will not be able to properly explore the parameter space.

4.1. Results

All Designs of Table 1 in total constitute for **13608** different Monte Carlo setups. We will present results only for those Designs that are sufficiently informative regarding general trends[19]. In particular, we do not present results for "intermediate designs" with $\theta = 0.3$, $\boldsymbol{\Upsilon}^{(5)}$. Furthermore, based on the preliminary MC results no substantial difference between two choices of $\boldsymbol{\Sigma}$ can be observed, so the results of $\boldsymbol{\Sigma}^{(2)}$ are omitted as well. We omit the results of $\alpha = -.9$ as they are both quantitatively and qualitatively similar to those of $\alpha = -.5$[20]. In total we are left with **2520** Designs with

[16] Circle is closed by connecting $i = 1$ with $i = N$.
[17] For a graphical illustration see Figure 2 of the aforementioned paper.
[18] Setting $S = 50$ would be another option, but it is of similar arbitrariness.
[19] All other results are available from the author upon request.
[20] The only difference is somewhat higher power at the costs of bigger size distortions for low N/moderate T.

summarized results presented in Tables A.3-A.6, henceforth Tables[21].

General Trends. First of all, we can observe that rejection frequencies are monotonically declining in $|\lambda|$ for the vast majority of Designs without spatial dependence (the case with spatial dependence will be discussed later). As we have discussed in Section 3.2 this property should not be taken as granted for the rk-J test (as dependence on $\boldsymbol{\Phi}$ in non-linear). It goes without saying that performance in terms of size and power is improving once N increases. However, for lower values of N test tends to be undersized for $T = 3$ and oversized for $T = 7$[22]. In the effect stationary case τ does not play substantial role and only affects the $\boldsymbol{V}$ matrix, but we can still observe that higher value of τ is associated with slightly lower power. The results of the effect non-stationary designs are way more interesting in this case and will be discussed afterwards. It is remarkable that for $N = 500$, the rk-J test has notable power when λ is very close to 0. For instance, all rejection frequencies in the effect stationary designs at $\lambda = 0.005$ are above 30% and 25% for $\alpha = -0.5$ and $\alpha = -0.1$ respectively.

As we have mentioned in the previous section, comparison between different values of α is not very *fair* and should be interpreted with caution (in the effect stationary case R^2_Δ for $\alpha = -0.5$ is roughly 5 times higher than the one for $\alpha = -0.1$.). In the vast majority of cases with size distortions being of similar magnitude test power for $\alpha = -0.5$ tends to be higher than for $\alpha = -0.1$ (with few exceptions when higher power is obtained at the costs of seriously oversized test 10%).

Spatial Dependence. Evidence of the uniform upward shift in the size can be observed when designs with spatial dependence ($\theta = 0.5$) are considered. This upward movement does not come as a surprise because similar patterns were documented in the panel unit root testing literature[23]. However, the same conclusion can not be reached regarding the test power, as for most scenarios it changes marginally and does not show any clear patterns in terms of magnitude and direction. More importantly, major size distortions *do* not disappear for $N = 500$, thus substantially higher cross-sectional dimension is required to perform inference with the correct size.

Effect Non-stationarity and Non-monotonic power curves. Firstly, we consider rejection frequencies for $\boldsymbol{\Upsilon} = 0.5\boldsymbol{I}_m$ as it is the most exceptional in terms of observed patters. In the latter case we observe power curves that are **not monotonic** for $\alpha = -0.1$ (especially for $N = 250$) and sharply decreasing for $\alpha = -0.5$ if $\tau = 5$ and $T = 3$. It can be intuitively explained as in this case the effect non-stationarity term in $\mathrm{E}[\Delta \boldsymbol{y}_{i,t}\boldsymbol{y}'_{i,t-1}]$ is *negative*, driving the whole expression towards zero matrix (recall analysis in Hayakawa (2009) for the univariate case). Thus, we have a weak instrument problem under alternative that is not induced by cointegration. These patterns are present irrespective of whether spatial dependence is present or not. As we have mentioned in Remark 4 in this MC study we do not fix any "orthogonal parameters" and that might be one of the reasons

[21]All rejection frequencies are rounded two digits after comma. Empty spots indicate maximal power. Numbers in **bold** indicate that the actual size is equal to the nominal one.

[22]As in this case orders of magnitude for N and T are not substantially different we suspect that critical values obtained as $N, T \to \infty$ (jointly) might be more appropriate.

[23]Moreover, presence of spatial dependence reduces the proportion (but not substantially) of var $(\boldsymbol{\varepsilon}_{i,t} + \boldsymbol{\eta}_i)$ attributed to the variance of the $\boldsymbol{\eta}_{i,t}$ that might substantially influence the results.

why *quadratic* power curves for $\tau = 5$ are observed[24]. By varying λ parameter we directly vary the relative contributions of fixed effects and idiosyncratic parts of the variance components in $\operatorname{var} \boldsymbol{y}_{i,t}$. For larger values of $|\lambda|$ the fixed effects part is more pronounced, resulting in substantial effects of the "negative" effect stationarity. On the other hand, for $|\lambda| \approx 0$ the idiosyncratic part is dominant and there is no substantial effects of the "negative" effect non-stationary initialization.

As it can be expected, the results for $\boldsymbol{\Upsilon} = 1.5\boldsymbol{I}_m$ are less controversial (again, recall Hayakawa (2009)). In this case power curves are monotonic, and rejection frequencies are uniformly dominating the ones from effect stationary case irrespective of other design parameters.
Results for $\boldsymbol{\Upsilon}^{(4)}$ seem to combine properties of both $\boldsymbol{\Upsilon}^{(3)}$ and $\boldsymbol{\Upsilon}^{(1)}$. From $\boldsymbol{\Upsilon}^{(1)}$ some non-monotonicities are inherited but for a smaller number of designs and are less pronounced. Apart from that, the superior test power properties (as compared to the effect stationary case) of $\boldsymbol{\Upsilon}^{(3)}$ are dominant. This combined behavior is due to the fact that $\boldsymbol{\Upsilon}^{(4)}$ is changing with λ. In designs with λ substantially lower than 0 we have $\boldsymbol{\Upsilon}^{(4)} \approx \boldsymbol{I}_m$, consecutively the *weak instrument problem under alternative* is less pronounced.
Finally, the results of $\boldsymbol{\Upsilon}^{(6)}$ are somewhat in between those of $\boldsymbol{\Upsilon}^{(1)}$ and $\boldsymbol{\Upsilon}^{(2)}$, but are slightly closer to $\boldsymbol{\Upsilon}^{(2)}$. It serves as an indication that the off-diagonal element in $\boldsymbol{\Upsilon}^{(6)}$ is not of any great importance (given the particular choice of other design parameters). It would be an interesting exercise to further investigate the impact of other choices of $\boldsymbol{\Upsilon}$ with possibly diagonal elements of different magnitude and sign. However, there are no clear guidelines for choosing *empirically* relevant $\boldsymbol{\Upsilon}$ in the literature.

5. Empirical Illustration

In this section we use the rk-J procedure to test for cointegration in Spanish firm panel dataset covering **1983-1990** of **738** manufacturing companies as in Alonso-Borrego and Arellano (1999). We construct bivariate PVAR(1) model where dependent variables are logs of employment and wages. It is reasonable to assume that time effects are present in the model so we explicitly consider variables in their deviations from the cross-sectional averages. Several alternative approaches for cointegration testing are considered. Firstly, we apply the rk test of KP directly to GMM estimates $\hat{\boldsymbol{\Pi}}$. We restrict set of GMM estimators to two step estimators (except for SGMM) presented in BHP (in total 3)[25]. Secondly, the LR tests based on the Transformed Maximum Likelihood function of BHP (LR-TMLE) and Conditional Maximum Likelihood function of Arellano (2003a) (LR-CMLE) are considered. Finally, the **rk-J** test of Section 3.2 is considered. Under $\mathbf{H}_0 : \operatorname{rk} \boldsymbol{\Pi} = 1$ all tests have limiting χ^2 distribution with one degree of freedom. Results are summarized in Table 2:
From Table 2 we can see that only the rk-J test based on the AH moment conditions rejects $\mathbf{H}_0$, while all the others do not reject it[26]. Numerous reasons might be responsible for differences in conclusions. First of all, we suspect that the initialization moment conditions of the System estimator are not valid and because of that it does not come as a surprise that all EGMM estimators fail to reject

[24] Some preliminary MC results, not presented in this paper suggest that effect of τ in this setup is not-monotonic. In the sense that higher values of τ will lead to increase of power rather than further decrease. At least for this particular design it seems that $\tau = 5$ represents the close to worst possible scenario as minimum is reached for $\tau \approx 6.2$.

[25] We use asymptotic standard errors for all GMM estimators. Here we do not present results based on Windmeijer (2005) type of finite sample corrected S.E., but they are available from the author upon request.

[26] We present results for **SGMM(2)** for informal comparison, as under $\mathbf{H}_0$ the AB (SGMM) estimator is not consistent.

Table 2: Cointegration testing. The 5% critical value is 3.84.

Name	Test Statistic
SGMM	12.271
EGMM1	0.89563
EGMM2	0.036158
EGMM3	2.0536
LR-TMLE	0.587821
LR-CMLE	0.546498
rk-J	13.352***

$\mathbf{H}_0$. Hayakawa and Nagata (2012) provide some evidence based on an incremental Sargan test in support of the latter statement[27]. Another explanation of results in Table 2 might be the low power of cointegration test used directly on the estimate of $\boldsymbol{\Pi}$. Some preliminary MC results suggest that at least for the System Estimator it might be the case, because the (size adjusted) power of the rk-J test dominates the power of rk-EGMM in most setups of Table 1.

Now we turn our attention to Likelihood Ratio tests. We know that both likelihood methods are robust to violations of mean stationarity, but are not so to time-series heteroscedasticity. Thus, we can not rule out the possibility that it can be one of the reasons for divergence in conclusions[28]. Furthermore, the possibility of potential model misspecification can not be ruled out, the rejection by rk-J test might indicate just that (recall that Arellano (2003b) considers PVAR(2), not the PVAR(1) model).

6. Discussions

In this section we will summarize the main theoretical problems and challenges for the rk-J test.
Effect non-stationarity. The importance of this problem can not be overestimated. As it has been already discussed both in Theoretical and in Monte Carlo sections, it is very difficult to draw general conclusions about the performance of the rk-J test when the $\{\boldsymbol{y}_{i,t}\}_{t=0}^{T}$ process is not effect stationary.

Initialization and common dynamics assumption. It was also assumed throughout the paper that the common dynamics assumptions is satisfied. In the univariate case we know that if this assumption is satisfied the moment conditions are not relevant at unity. On the other hand, if the common dynamics assumptions is violated it is possible to have non-zero correlation at the unit root, see an extensive overview in Bun and Kleibergen (2012) on this issue.

Cross-sectional heterogeneity. Another issue reserved for future investigation in this paper is cross-sectional heterogeneity in $\boldsymbol{\Pi}$ parameter matrix. If $\boldsymbol{\Pi}_i$ matrices are individually specific then it is possible that individually all cross-sectional units have cointegrated dynamics but the average is not (or the other way around). Hence, in situations with possible cross-sectional heterogeneity in $\boldsymbol{\Pi}_i$ we should be very cautious interpreting the rejection of $\mathbf{H}_0$.

[27] However, this testing procedure can not be used if series are cointegrated.

[28] Arellano (2003b) presents some evidence of time-series heteroscedasticity in this dataset.

7. Conclusions

In this paper we have studied the properties of the standard Anderson and Hsiao (1982) moment conditions in a PVAR(1) when the process is cointegrated. Under the assumptions similar to Binder et al. (2005) we have shown that these moment conditions are of reduced rank if and only if the process is cointegrated. Based on this observation we have proposed a rank based test to test the null hypothesis of cointegration. The suggested testing procedure was shown to be robust to both cross-sectional and time series heteroscedasticity as well as spatial dependence. In the Monte Carlo study we found evidence that the test is reasonably sized and has good power properties in the vast majority of scenarios but might exhibit non-monotonic power curves for models with substantial effect non-stationarity. We have applied our testing procedure to the Spanish manufacturing data of Arellano (2003b) and unlike the test of BHP we have rejected the null hypothesis of cointegration.

References

Alonso-Borrego, C. and M. Arellano (1999). Symmetrically normalized instrumental-variable estimation using panel data. *Journal of Business & Economic Statistics 17*, 36–49.

Anderson, T. W. and C. Hsiao (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics 18*, 47–82.

Anderson, T. W. and H. Rubin (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics 20*, 46–63.

Anselin, L. (2003). Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review 26*, 153–166.

Arellano, M. (2003a). Modeling optimal instrumental variables for dynamic panel data models. Unpublished manuscript.

Arellano, M. (2003b). *Panel Data Econometrics*. Advanced Texts in Econometrics. Oxford University Press.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies 58*, 277–297.

Bai, J. and S. Ng (2004). A panic attack on unit roots and cointegration. *Econometrica 72*, 1127–1177.

Baltagi, B. H., G. Bresson, and A. Pirotte (2007). Panel unit root tests and spatial dependence. *Journal of Applied Econometrics 22*, 339–360.

Binder, M., C. Hsiao, and M. H. Pesaran (2005). Estimation and inference in short panel vector autoregressions with unit root and cointegration. *Econometric Theory 21*, 795–837.

Breitung, J. and H. M. Pesaran (2008). Unit roots and cointegration in panels. In L. Mátyás and P. Sevestre (Eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Advanced Studies in Theoretical and Applied Econometrics, Chapter 9, pp. 279–322. Springer Verlag.

Bun, M. J. G. and F. Kleibergen (2012). Identification in linear dynamic panel data models. Working Paper.

Hayakawa, K. (2009). On the effect of mean-nonstationarity in dynamic panel data models. *Journal of Econometrics 153*, 133–135.

Hayakawa, K. (2011, February). An improved gmm estimation of panel var models with applications to granger causality test and impulse response analysis. Working paper.

Hayakawa, K. and S. Nagata (2012). On the behavior of the gmm estimator in persistent dynamic panel data models with unrestricted initial conditions. Working Paper.

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

Kelejian, H. H. and I. R. Prucha (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroscedastict disturbances. *Journal of Econometrics 157*, 53–67.

Kiviet, J. F. (2007). Judging contending estimators by simulation: Tournaments in dynamic panel data models. In G. Phillips and E. Tzavalis (Eds.), *The Refinement of Econometric Estimation and Test Procedures*, Chapter 11, pp. 282–318. Cambridge University Press.

Kiviet, J. F. (2012). Monte carlo simulation for econometricians. In W. H. Greene (Ed.), *Foundations and Trends in Econometrics*, Volume 5. NOW the essence of knowledge.

Kleibergen, F. (2005). Testing parameters in gmm without assuming that they are identified. *Econometrica 73*(4), 1103–1123.

Kleibergen, F. and S. Mavroeidis (2009). Inference on subsets of parameters in gmm without assuming identification. Working Paper.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*, 97–126.

Lütkepohl, H. (2006). *New Introduction to Multiple Time Series Analysis* (2nd ed.). Springer.

Magnus, J. R. and H. Neudecker (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons.

Mutl, J. (2009). Panel var models with spatial dependence. Working Paper.

Pesaran, M. H. and E. Tosseti (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics 161*, 182–202.

Phillips, P. C. B. (1989). Partially identified econometric models. *Econometric Theory 5*, 181–240.

Robertson, D., V. Sarafidis, and J. Symons (2010). Iv estimation of panels with factor residuals. Working Paper.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics 71*, 135–158.

Sarafidis, V. (2011). Gmm estimation of short dynamic panel data models with error cross-sectional dependence. working paper.

Sarafidis, V. and T. Wansbeek (2012). Cross-sectional dependence in panel data analysis. *Econometric Reviews 31*(5), 483–531.

Stock, J. and J. Wright (2000). Gmm with weak identification. *Econometrica 68*(5), 1055–1096.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics 126*(1), 25–51.

Appendices

Appendix A. Tables and Graphs

Table A.3: $\alpha = -0.1$. $\boldsymbol{\Upsilon}^{(1)-(3)}$

		$\boldsymbol{\Upsilon}^{(1)}$							$\boldsymbol{\Upsilon}^{(2)}$							$\boldsymbol{\Upsilon}^{(3)}$						
θ	$\{N,T,\tau\} \setminus \lambda$	-.70	-.30	-.10	-.05	-.01	-.005	.00	-.70	-.30	-.10	-.05	-.01	-.005	.00	-.70	-.30	-.10	-.05	-.01	-.005	.00
0.0	50,3,1	.83	.81	.63	.36	.07	.04	.03	.90	.86	.64	.37	.07	.04	.03	.94	.90	.71	.42	.08	.05	.03
	50,3,5	.20	.20	.21	.14	.04	.03	.02	.62	.57	.37	.21	.05	.04	.03	.94	.88	.69	.43	.09	.05	.03
	50,5,1	.98	.95	.91	.75	.16	.08	.04		.99	.93	.76	.17	.08	.04		.99	.96	.81	.19	.09	.04
	50,5,5	.68	.40	.40	.33	.08	.04	.03	.88	.83	.69	.51	.12	.07	.03		.99	.93	.79	.20	.10	.04
	50,7,1		.99	.97	.92	.29	.12	**.05**			.99	.93	.30	.13	**.05**				.95	.31	.14	**.05**
	50,7,5	.97	.63	.52	.48	.15	.07	.04	.95	.92	.84	.71	.22	.11	**.05**			.98	.93	.31	.14	**.05**
	250,3,1	.99				.41	.16	.04				.99	.41	.16	.04					.44	.17	.04
	250,3,5	.21	.26	.65	.66	.17	.06	.02			.98	.90	.30	.11	.03				.99	.44	.17	.04
	250,5,1					.77	.34	.04					.77	.34	**.05**					.78	.35	**.05**
	250,5,5	.97	.65	.80	.85	.46	.17	.03					.65	.28	.04					.76	.34	**.05**
	250,7,1					.92	.51	**.05**					.92	.51	**.05**					.93	.52	**.05**
	250,7,5		.94	.88	.90	.66	.29	.04					.84	.43	**.05**					.91	.49	**.05**
	500,3,1					.75	.32	.04					.75	.33	.04					.77	.34	**.05**
	500,3,5	.24	.31	.85	.91	.45	.16	.03					.62	.26	.04					.75	.33	**.05**
	500,5,1					.97	.62	**.05**					.97	.61	**.05**					.97	.62	**.05**
	500,5,5		.80	.93	.96	.79	.39	.03					.93	.53	**.05**					.97	.60	**.05**
	500,7,1						.80	**.05**						.80	**.05**						.81	**.05**
	500,7,5		.99	.97	.98	.90	.58	**.05**					.99	.72	**.05**						.78	**.05**
0.5	50,3,1	.84	.82	.63	.37	.09	.07	**.05**	.89	.85	.64	.37	.09	.06	**.05**	.93	.88	.69	.41	.10	.07	**.05**
	50,3,5	.27	.27	.27	.18	.06	.05	.04	.67	.61	.40	.24	.07	.05	.04	.93	.86	.67	.41	.10	.07	**.05**
	50,5,1	.99	.97	.91	.74	.20	.11	.07	.99	.99	.93	.75	.20	.11	.07		.99	.95	.79	.21	.12	.07
	50,5,5	.75	.50	.49	.39	.11	.07	**.05**	.91	.86	.73	.53	.15	.09	.06		.99	.93	.77	.21	.12	.07
	50,7,1		.99	.98	.91	.31	.16	.08			.99	.92	.32	.16	.08			.99	.94	.33	.17	.08
	50,7,5	.98	.72	.63	.56	.19	.11	.07	.97	.94	.87	.75	.25	.14	.07			.98	.92	.32	.17	.08
	250,3,1				.99	.42	.19	.06				.99	.42	.19	.07				.99	.44	.20	.07
	250,3,5	.40	.47	.80	.78	.24	.11	.04			.98	.92	.33	.15	**.05**				.99	.43	.20	.07
	250,5,1					.75	.36	.08					.74	.36	.08					.75	.37	.08
	250,5,5	.98	.81	.91	.93	.55	.23	**.05**					.65	.31	.07					.74	.36	.08
	250,7,1					.90	.52	.09					.90	.51	.09					.91	.52	.08
	250,7,5		.97	.95	.96	.74	.38	.07					.84	.45	.08					.89	.50	.08
	500,3,1					.72	.34	.08					.72	.34	.08					.73	.35	.08
	500,3,5	.50	.61	.95	.97	.54	.22	.06					.63	.28	.07					.72	.34	.07
	500,5,1					.95	.61	.08					.95	.60	.08					.96	.61	.08
	500,5,5		.92	.98	.99	.86	.47	.07					.92	.54	.08					.95	.60	.08
	500,7,1						.77	.08						.77	.08						.78	.08
	500,7,5			.99	.99	.96	.65	.07					.99	.71	.08					.99	.76	.08

Table A.4: $\alpha = -0.1$. $\boldsymbol{\Upsilon}^{(4),(6)}$

		$\boldsymbol{\Upsilon}^{(4)}$							$\boldsymbol{\Upsilon}^{(6)}$						
θ	$\{N,T,\tau\} \setminus \lambda$	-.70	-.30	-.10	-.05	-.01	-.005	.00	-.70	-.30	-.10	-.05	-.01	-.005	.00
0.0	50,3,1	.92	.89	.66	.38	.07	.04	.03	.88	.84	.64	.37	.07	.04	.03
	50,3,5	.79	.81	.52	.21	.04	.03	.03	.36	.39	.30	.19	.05	.03	.02
	50,5,1		.99	.94	.76	.17	.08	.04	.99	.97	.92	.76	.17	.08	.04
	50,5,5	.97	.98	.85	.50	.09	.05	.03	.66	.59	.56	.45	.11	.06	.03
	50,7,1			.99	.93	.29	.12	**.05**		.99	.98	.92	.30	.13	**.05**
	50,7,5	.99		.95	.71	.17	.09	.04	.84	.73	.70	.63	.21	.10	.04
	250,3,1					.43	.16	.04				.99	.41	.16	.04
	250,3,5				.94	.22	.08	.03	.74	.85	.92	.85	.29	.11	.03
	250,5,1					.79	.35	**.05**					.77	.34	**.05**
	250,5,5				.99	.52	.19	.04	.95	.93	.97	.97	.63	.27	.04
	250,7,1					.93	.52	**.05**					.92	.51	**.05**
	250,7,5					.71	.32	.04	.99	.97	.99	.99	.83	.42	**.05**
	500,3,1					.77	.33	.04					.75	.32	.04
	500,3,5					.50	.18	.03	.89	.96	.99	.99	.62	.25	.04
	500,5,1					.98	.63	**.05**					.97	.61	**.05**
	500,5,5					.81	.40	.04	.99	.99			.92	.52	.04
	500,7,1						.81	**.05**						.80	**.05**
	500,7,5					.92	.60	**.05**					.99	.72	**.05**
0.5	50,3,1	.87	.84	.63	.37	.10	.06	**.05**	.87	.84	.63	.37	.10	.06	**.05**
	50,3,5	.44	.45	.35	.22	.07	.05	.04	.44	.45	.35	.22	.07	.05	.04
	50,5,1	.99	.98	.92	.74	.20	.11	.07	.99	.98	.92	.74	.20	.11	.07
	50,5,5	.75	.68	.63	.49	.15	.09	.06	.75	.68	.63	.49	.15	.09	.06
	50,7,1			.98	.92	.32	.16	.08			.98	.92	.32	.16	.08
	50,7,5	.89	.81	.77	.68	.24	.13	.07	.89	.81	.77	.68	.24	.13	.07
	250,3,1				.99	.42	.19	.07				.99	.42	.19	.07
	250,3,5	.87	.93	.95	.88	.32	.15	**.05**	.87	.93	.95	.88	.32	.15	**.05**
	250,5,1					.74	.36	.08					.74	.36	.08
	250,5,5	.98	.97	.99	.99	.64	.30	.07	.98	.97	.99	.99	.64	.30	.07
	250,7,1					.90	.51	.09					.90	.51	.09
	250,7,5		.99	.99		.83	.45	.08		.99	.99		.83	.45	.08
	500,3,1					.72	.34	.08					.72	.34	.08
	500,3,5	.97	.99		.99	.62	.28	.06	.97	.99		.99	.62	.28	.06
	500,5,1					.95	.60	.08					.95	.60	.08
	500,5,5					.91	.54	.08					.91	.54	.08
	500,7,1						.77	.08						.77	.08
	500,7,5					.98	.71	.08					.98	.71	.08

Table A.5: $\alpha = -0.5$. $\boldsymbol{\Upsilon}^{(1)-(3)}$

		$\boldsymbol{\Upsilon}^{(1)}$							$\boldsymbol{\Upsilon}^{(2)}$							$\boldsymbol{\Upsilon}^{(3)}$						
θ	$\{N,T,\tau\} \setminus \lambda$	-.70	-.30	-.10	-.05	-.01	-.005	.00	-.70	-.30	-.10	-.05	-.01	-.005	.00	-.70	-.30	-.10	-.05	-.01	-.005	.00
0.0	50,3,1	.99	.88	.59	.36	.08	.05	.04	.99	.97	.72	.45	.10	.06	.04			.84	.54	.12	.07	.03
	50,3,5	.94	.34	.09	.08	.03	.02	.02	.86	.64	.36	.23	.08	.05	.03	.99	.97	.83	.54	.13	.08	.04
	50,5,1		.99	.91	.79	.25	.13	.06			.97	.85	.29	.15	.07			.99	.89	.30	.16	.06
	50,5,5		.88	.26	.21	.08	.05	.03	.97	.89	.74	.59	.22	.13	.06	.97	.99	.96	.85	.27	.14	.06
	50,7,1			.98	.94	.43	.23	.10				.97	.45	.24	.10				.98	.45	.23	.09
	50,7,5		.99	.54	.40	.18	.10	.06		.97	.90	.80	.37	.21	.09	.95	.99	.98	.95	.40	.21	.07
	250,3,1			.99	.98	.42	.16	.03				.99	.47	.20	.04					.49	.22	**.05**
	250,3,5		.95	.09	.14	.08	.04	.02			.95	.87	.36	.16	.04					.47	.21	**.05**
	250,5,1					.82	.42	.06					.82	.42	.06					.82	.42	**.05**
	250,5,5			.56	.24	.13	.06	.02				.99	.72	.37	.06					.78	.39	**.05**
	250,7,1					.95	.60	.07					.94	.60	.07					.95	.60	.06
	250,7,5			.92	.66	.40	.22	.04					.89	.54	.07					.93	.55	.06
	500,3,1					.76	.35	.04					.78	.37	**.05**					.79	.38	**.05**
	500,3,5			.09	.20	.20	.08	.02				.99	.67	.31	.04					.76	.36	**.05**
	500,5,1					.98	.68	.06					.98	.68	**.05**					.98	.68	**.05**
	500,5,5			.79	.23	.14	.07	.03					.94	.61	**.05**					.97	.64	**.05**
	500,7,1						.85	.06						.85	.06						.85	.06
	500,7,5			.99	.79	.48	.30	.04					.99	.79	.06						.82	**.05**
0.5	50,3,1	.99	.92	.64	.40	.12	.08	.06		.97	.73	.47	.14	.09	.06			.83	.54	.15	.09	.06
	50,3,5	.94	.39	.12	.10	.04	.04	.03	.90	.71	.41	.27	.10	.07	**.05**		.98	.81	.53	.15	.10	.06
	50,5,1		.99	.93	.80	.29	.17	.10			.98	.84	.31	.19	.11			.99	.87	.32	.19	.11
	50,5,5		.89	.35	.28	.11	.07	**.05**	.98	.93	.78	.63	.26	.16	.10	.98	.99	.97	.84	.30	.17	.09
	50,7,1			.99	.95	.45	.27	.14				.96	.46	.27	.14				.97	.45	.27	.13
	50,7,5		.99	.64	.51	.23	.14	.09		.98	.93	.83	.40	.25	.13	.97		.99	.95	.42	.24	.10
	250,3,1				.98	.45	.21	.06				.99	.48	.23	.07					.49	.24	.08
	250,3,5		.93	.17	.14	.06	.03	.02			.97	.89	.38	.19	.06				.99	.46	.23	.08
	250,5,1					.79	.43	.09					.79	.44	.09					.79	.44	.09
	250,5,5			.68	.45	.23	.11	.03					.72	.39	.09					.76	.40	.08
	250,7,1					.92	.59	.11					.92	.59	.11					.93	.59	.10
	250,7,5			.95	.83	.59	.34	.07					.88	.55	.10					.91	.55	.09
	500,3,1					.74	.38	.07					.75	.38	.08					.76	.39	.08
	500,3,5			.20	.14	.10	.05	.02					.67	.34	.07					.74	.37	.07
	500,5,1					.96	.66	.09					.96	.66	.09					.96	.66	.09
	500,5,5			.86	.53	.31	.16	.03					.94	.61	.08					.96	.63	.08
	500,7,1						.83	.10						.82	.10						.83	.09
	500,7,5			.99	.93	.77	.53	.07					.99	.78	.09						.80	.08

Table A.6: $\alpha = -0.5$. $\boldsymbol{\Upsilon}^{(4),(6)}$

		$\boldsymbol{\Upsilon}^{(4)}$							$\boldsymbol{\Upsilon}^{(6)}$						
θ	$\{N,T,\tau\} \setminus \lambda$	-.70	-.30	-.10	-.05	-.01	-.005	.00	-.70	-.30	-.10	-.05	-.01	-.005	.00
0.0	50,3,1		.95	.74	.46	.11	.06	.04	.98	.92	.68	.43	.10	.06	.04
	50,3,5		.50	.45	.26	.09	.05	.04	.58	.30	.22	.17	.07	.04	.03
	50,5,1		.99	.98	.86	.30	.15	.07		.98	.95	.83	.28	.15	.07
	50,5,5		.68	.82	.63	.27	.14	.07	.92	.57	.44	.43	.19	.11	**.05**
	50,7,1				.97	.47	.25	.10			.99	.95	.45	.24	.10
	50,7,5		.84	.94	.81	.46	.25	.10		.82	.62	.62	.34	.20	.08
	250,3,1				.99	.50	.21	.04				.99	.47	.20	.04
	250,3,5		.95	.98	.89	.47	.20	.04	.91	.59	.58	.60	.28	.13	.03
	250,5,1					.84	.44	.06					.82	.42	**.05**
	250,5,5		.97		.99	.84	.47	.06		.89	.81	.88	.67	.35	**.05**
	250,7,1					.96	.63	.07					.94	.60	.07
	250,7,5		.99			.95	.66	.07		.99	.92	.96	.87	.54	.06
	500,3,1					.81	.39	**.05**					.78	.37	**.05**
	500,3,5		.99		.99	.80	.39	.04	.98	.78	.81	.85	.57	.27	.03
	500,5,1					.98	.71	**.05**					.98	.68	**.05**
	500,5,5					.99	.73	**.05**		.97	.94	.98	.93	.60	**.05**
	500,7,1						.87	.06						.85	.06
	500,7,5						.88	.06			.98		.99	.80	.06
0.5	50,3,1		.96	.74	.49	.14	.10	.07	.99	.94	.70	.46	.14	.09	.06
	50,3,5		.58	.48	.30	.11	.08	**.05**	.68	.39	.28	.21	.09	.06	.04
	50,5,1			.98	.85	.33	.19	.11		.99	.96	.83	.31	.19	.11
	50,5,5		.76	.85	.66	.30	.18	.10	.95	.67	.54	.50	.23	.15	.08
	50,7,1				.96	.47	.28	.14			.99	.95	.46	.27	.14
	50,7,5		.90	.96	.84	.47	.28	.14		.88	.71	.69	.37	.23	.12
	250,3,1				.99	.50	.24	.07				.99	.48	.23	.07
	250,3,5		.97	.99	.91	.47	.23	.06	.96	.78	.73	.71	.33	.17	**.05**
	250,5,1					.81	.45	.10					.79	.44	.09
	250,5,5		.99		.99	.81	.46	.09		.95	.92	.95	.70	.38	.09
	250,7,1					.93	.60	.10					.92	.59	.11
	250,7,5					.94	.63	.10			.97	.99	.87	.55	.10
	500,3,1					.77	.41	.08					.75	.39	.08
	500,3,5					.77	.40	.07	.99	.93	.92	.93	.62	.31	.06
	500,5,1					.97	.68	.09					.96	.66	.09
	500,5,5					.97	.70	.08		.99	.98	.99	.93	.61	.08
	500,7,1						.84	.10						.83	.10
	500,7,5						.85	.09					.99	.79	.09