Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data*

Anindya Banerjee[†] Massimiliano Marcellino[‡] Chiara Osbat[§]

This version: July 2000

Abstract

We show how the use of panel data methods such as those proposed in single equations by Kao (1999) and Pedroni (1999) or in systems by Larsson and Lyhagen (1999) to investigate economic hypotheses such as purchasing power parity or the term structure of interest rates may be affected by the existence of cross-unit cointegrating relations. The existing literature assumes that such relations, that tie the units of the panel together, are not present. Using empirical examples from a panel of OECD countries we show that this assumption is very likely to be violated. Simulations of the properties of panel cointegration tests in the presence of cross-unit relations are then presented to demonstrate the serious cost of assuming away such relations. Some fixes are proposed as a way of dealing with these more general scenarios.

*We are very grateful to the European University Institute for partially funding this research under research project "Modelling and Forecasting Economic Time Series in the Presence of High Orders of Integration and Structural Breaks in Data".

[†]Corresponding author. Department of Economics, European University Institute, 50016 San Domenico di Fiesole. Phone: +39-055-4685-356. Fax: +39-055-4685-202. E-mail: banerjee@datacomm.iue.it

[‡]Istituto di Economia Politica and IGIER, Universita' Bocconi, Via Salasco 3, 20136 Milano

[§]Department of Economics, European University Institute, 50016 San Domenico di Fiesole

1 Introduction

In recent years, the use of panel data techniques to test macroeconomic hypotheses has become increasingly common. Two examples of the most popular uses to which these techniques have been put are testing for convergence in growth rates of GDP[1] and tests for purchasing power parity[2]. Each of these hypotheses is formulated essentially as testing for a unit root in the estimated residuals of a single-equation regression and the argument presented in favour of panel data methods is that they have greater power than standard unit root tests, by virtue of the reduction in noise caused by the averaging or pooling across the units of the panel. An added advantage often proposed is that the asymptotic distribution (under the null) of some of these tests, even in the presence of integrated data, is normal. Inference is thereby in some sense made "standard". The emphasis in this literature has therefore primarily taken the form of considering the asymptotic properties of panel data estimators and test statistics as T and N go to infinity, possibly at rates of divergence that are controlled appropriately. Techniques developed by McKoskey and Kao (1998), Kao (1999) and Pedroni (1999) *inter alia* have extended Engle (1987) type static single equation regressions to pooled panel static regressions and produced test statistics and critical values to test for cointegration in panel data. Implicit in these analyses is the assumption of unique cointegrating vectors, albeit heterogeneous, across the units.

Several issues are worth noting within the context of this study for the use of panels in macro-econometric research. It is important firstly to make a clear distinction between their use here and their use in micro-econometric research. The macroeconomic data sets (relating to growth, distribution, term structure of interest rates or PPP) typically consist of samples of up to 50 countries with observations

[1]See e.g. Evans and Karras (1996) and Lee, Pesaran, and Smith (1997)

[2]On panel studies of PPP, see Frankel and Rose (1996) and Papell (1997).

for thirty or forty years quarterly or annually. Thus the dimension of the number of cross-section units N, although sometimes quite large, is very often dominated by the time dimension T. The uses of panel data methods in micro-econometrics on the other hand has a much longer history and involve looking at settings where the time dimension T is significantly smaller than N. Thus, while an impressive amount of research has focused on modelling the heterogeneity across the units by means of common time effects, or fixed- and random-effects, and on methods of overcoming what is commonly known as the "incidental parameters problem" (usually by means of transformations of the model or by instrumental variables or GMM methods), the issue is dealt with by considering large-T, relatively-large-N samples in the case of macro-econometric data sets. This is not to say that incidental parameter problems do not have any relevance for our analysis, since the assumption that we have sufficient data to estimate all the parameters of our model is far from innocuous and we shall visit this issue below.

Secondly, in the context of analysis of multivariate (> 2) data sets it is natural to focus on methods that relax the assumption of a unique cointegrating vector. Groen and Kleibergen (1999), Larsson and Lyhagen (1999) and (2000), and Larsson, Lyhagen, and Löthgren (1998) have therefore developed techniques analogous to (Johansen 1995) maximum likelihood methods to allow for multiple cointegrating vectors in the cross-section units. Since the assumption of unique cointegrating vectors may be thought of as being unnecessarily restrictive and unrealistic, systems-methods for panels have represented a notable advance.

However, thirdly, and most importantly, none of the methods cited above takes any account of the possibility of long-run dependence among the variables across the units comprising the panel (except through short-run effects or through correlations among the errors across the units). As we show below, particularly in macroeconomic studies with integrated variables, this is a serious restriction. While GLS methods such as those used by OConnell (1998) go some way towards allowing the units to be related by means of a non-diagonal variance-covariance matrix of the system, these do not account for the existence of long-run or cointegrating relationships across variables in the units of the panel. Nor is the incorporation of short-run dependence sufficient for the analysis of macroeconomic panel data sets.

This last observation leads on to important complications for the research agenda involving the use of macroeconomic panel data. Within the compass of this critique lie all the pre-existing tests for unit roots and cointegration in panels, to the extent of vitiating entirely their use. Two significant features inhibit their applicability - the assumptions of the same maximum cointegrating rank across the units and no long-run dependence across the units.

Our paper takes Larsson and Lyhagen (1999) (henceforth LL) as its starting point, since it provides one of the most general treatments of panel cointegration (subject to the restrictions highlighted above). The motivation for and background to our study is the consideration of small- to medium-sized macro-economic panels, with the time dimension typically ranging from 100 to 200. Our methodology is based on simulations. In Section 2, we briefly present the maximum likelihood framework for analyzing panel data as exemplified by LL who also derive the null distributions of the test statistics. In Section 3, we use an empirical example with two units in the panel to illustrate the considerable difficulties for the direct use of this method. By extension, any of the single-equation methods cited above may be shown to be in similar trouble. By means of simple studies of long- and short-run interest rates in a small panel where the cross-country cointegration possibilities are self-evident, we motivate strongly the need to modify pre-existing methods of looking at cointegrated panels.

In order to bench-mark our simulation study, in Section 4 we simulate data under the LL specifications of the data generation process and show that use of the asymptotic distributions and critical values derived by LL leads to tests with correct size and good power. Our point of departure from this framework is then to simulate data with cross-unit cointegrating relations or units with different cointegrating rank and to show the properties of the LL tests in their presence. Notably, in Sections 5 and 6 we present results demonstrating findings of incorrect rank under simple changes of the process generating the data.

We also suggest and evaluate modifications of the testing framework to allow for cross-unit cointegration and to improve inference. An obvious route we explore is to pre-test the units of the panel for cointegrating relations within each unit using the Johansen ML method. The LL method is applicable if and only if one is entitled to allow for the same maximum cointegrating rank across the units of the sample and, moreover, cointegration occurs only within units and not across units. Estimating the system unit by unit using Johansen ML is a way of verifying the first assumption. If this is verified, one may then extract the common trends implied by the cointegrating vectors and test for cointegration among the common trends *à la* Gonzalo and Granger (1995) to rule out the existence of cross-unit relations. Only then is the LL approach (and the special cases of this approach, such as Kao (1999) and Pedroni (1999)) justified, and leads to gains in efficiency over estimating the full system (consisting of all the variables in all the units of the panel). This sequential testing procedure is implemented in Section 7 in order to analyze in more detail the interest rate data. The analysis clearly indicates that the panel cointegration techniques are not suited to this context, and can lead to incorrect conclusions.

The limitations of our current study must of course be borne in mind. Our investigation is limited to looking at only the simplest possible scenarios. We do not as yet have sufficient evidence for larger dimensional systems. Nor do we present here the results for data generation processes with deterministic components such as constants or trends or break dummies, complications that are very likely to occur in practice. In this latter case, provided these deterministic variables enter into the system unrestrictedly, we may consider our analysis as proxying for as if we were working with detrended or demeaned data. Nevertheless, based on the consideration of even these simple cases, we are drawn inexorably to the conclusion that in general only panel methods that allow for full system maximum likelihood analysis are likely to lead to the "right" answer. However, data limitations, with the implied degrees of freedom restriction, will not allow such a full system to be estimated unrestrictedly in all circumstances. We therefore conclude that, particularly in integrated panels, great care must be taken in using and interpreting the results of macroeconomic panels. Short-cuts or restrictions of the kind implicit in LL, or fixes such as those which would enable the within-unit cointegrating relations to be estimated first and then imposed in order to look at the cross-unit cointegrating relations, could in some circumstances be useful but prove to be profoundly misleading in others.

2 The Panel Cointegration Framework

The LL panel cointegration model corresponds to a restricted cointegrated VAR. The model under consideration is

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{k=1}^{m-1} \Gamma_k \Delta Y_{t-k} + \varepsilon_t$$

where α and β have dimension $Np \times \sum_{i=1}^{N} r_i$, N is the number of units, p is the dimension of each sub-system $Y_{it} = (y_{i1t}, y_{i2t}, ..., y_{ipt})', i = 1, 2, ...N,\ t = 1, 2, ..., T,$

and r_i is the rank of each sub-system. The vector Y_t is given by stacking the N vectors Y_{it}. The matrices α and β have the following structure:

$$\alpha = \begin{pmatrix} \alpha_{11} & \alpha_{22} & \cdots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & & \\ \vdots & & \ddots & \vdots \\ \alpha_{N1} & \alpha_{N2} & \cdots & \alpha_{NN} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_{11} & 0 & \cdots & 0 \\ 0 & \beta_{22} & & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_{NN} \end{pmatrix}.$$

The variance-covariance matrix of ε_t is

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1N} \\ \Sigma_{21} & \Sigma_{22} & .... & \Sigma_{2N} \\ \vdots & & \ddots & \vdots \\ \Sigma_{N1} & \Sigma_{N2} & \cdots & \Sigma_{NN} \end{pmatrix}$$

The model therefore allows for interaction among the units through the long-run adjustment coefficients α, the short-run matrices Γ_k, and the off-diagonal elements of Σ, but the restriction $\beta_{ij} = 0 \; \forall i \neq j$ rules out cointegrating relationships across the units. A further restriction is that of r_i being assumed to be the same for each unit. These two restrictions are the most important issues to be investigated, since there are many examples where they may be inconsistent with both theory and data.[3]

The estimation algorithm for the individual cointegrating relations is a series of reduced rank regressions in which each β_{ii} is estimated by concentrating out the $n-1$ remaining ones. Hence, β_{11} to β_{NN} are estimated at each iteration, and the

[3]Note that in practice, as N grows, the amount of correlation among the units must be controlled *a priori* by imposing restrictions on the α, Γ and Σ matrices in order to satisfy degrees of freedom restrictions implied by the data.

procedure is repeated until convergence.[4]

The trace test for cointegrating rank of $r_i = q$ *versus* $r_i = p$, for each i and for $q = 0, ..., p-1$, is derived in LL. The rank is tested under the block-diagonality restriction on β. The asymptotic distribution of the test statistic is the convolution of the standard Johansen LR rank test and an independent χ^2 with $N(N-1)(p-q)q$ degrees of freedom. This means that the test for $r_i = 0$ is the same as the Johansen test, while for rank larger than zero there is an additional component in the distribution which accounts for the additional zero restrictions imposed on β. Note that testing for $r_i = q$ *versus* $r_i = p$ in this framework corresponds to testing for rank Nq *versus* Np in a full system context, whereas testing for rank in the Johansen framework would allow for all the intermediate possibilities. For example, if $N = p = 2$, LL tests for 0 *versus* 4 followed by 2 *versus* 4. Johansen would test for 0 *versus* 4, 1 versus 4 and so on.

Finally, note that when $p = 1$ and Γ_k is block diagonal for each k, the null hypothesis of the LL test is that all the variables are I(1), against the alternative that they are stationary with heterogenous stationary roots. This result is therefore the panel version of the ADF test in the single unit case, and in this sense the LL framework nests panel unit root tests (see *e.g.* Levin and Lin (1993), Im and M. H.and Shin (1997)).

3 Empirical Examples

To illustrate the kind of problem one could encounter in using the LL panel cointegration model in empirical modelling, we present two simple examples involving short- and long-term interest rates.[5] We estimate the two systems consisting of Ger-

[4]For starting values, LL propose using the β_{ii} estimated from a standard cointegration analysis on each unit separately. We instead use the initial values suggested in (Johansen 1995), p. 110.

[5]The variables, taken from the OECD database, are three-month interest rates (code 6225D) and long term rates (codes 6253D for Germany, 6261D for Denmark and 6269D for the Netherlands).

many and the Netherlands, and Germany and Denmark by full system maximum likelihood, using the Johansen approach. It turns out that both pairs of countries exhibit a cointegration pattern that is not consistent with the structure of the LL model.[6]

3.1 Germany and The Netherlands

The sample period for the German-Dutch model goes from May 1990 to January 1999. The chosen specification for the unrestricted VAR involves four lags and an unrestricted constant. Diagnostic tests for this simple model point towards its congruence. There is only a marginal rejection of the null hypothesis of normality in the system, which is not likely to affect the outcome of the cointegration tests, see e.g. Gonzalo (1994).

The Johansen trace and maximum eigenvalue cointegration test statistics in this model indicate that the cointegrating rank is equal to one, see the first panel of Table 1. This is an indication that using the LL model would be inappropriate.

The restricted cointegration vector corresponding to a rank of one involves short-term rates from both countries, with homogeneous coefficients equal to one. Only the German short-term interest rate adjusts to disequilibrium. The LR test corresponding to this restricted specification is distributed as a $\chi^2_{(6)}$ and takes a value of 4.84, corresponding to a tail probability value of 0.56. The estimated coefficients are reported in the second panel of Table 1. It is evident that with this cointegration structure it would be inappropriate to estimate the LL panel cointegration model.

This part of the analysis was conducted using PcFiml 9.21 Doornik and Hendry (1997)

[6]We are very grateful to Johan Lyhagen for providing us with the Larsson and Lyhagen panel data programs written in GAUSS 3.0. This code was subsequently translated by us to OX (Doornik (1999)) and modified and extended for the uses made in this paper and is freely available from us.

3.2 Germany and Denmark

For the German-Danish model, the sample period considered goes from May 1990 to January 1999. We adopt the same specification as above, with four lags and an unrestricted constant. In this sample, there is evidence of parameter instability in 1993 in the Danish short-term rate and in the German long-term rate. However, the only evidence of misspecification is the rejection of normality in the Danish short- and long-term rates and in the full system. If we analyze this model on the sample starting from October 1993, we find no evidence of parameter instability. The cointegration analysis suggests a rank of one in both samples, and we report the results for the longer sample only.

The Johansen cointegration tests, reported in the first panel of Table 2, indicate that the rank is one in this bivariate system as well.

The restricted cointegrating vector involves the short-term rates only, and the German long-term rate is weakly exogenous. The LR test for this specification is distributed as a $\chi^2_{(3)}$ and takes the value of 4.92, corresponding to a tail probability of 0.18. Parameter estimates for the restricted model are reported in the second panel of Table 2.

In summary, in both cases the block diagonality of β that characterizes the panel cointegration framework is violated. We will evaluate the consequences of this violation on the performance of the LL test in Section 5, while the next section examines the size and power of the test, assuming that its underlying assumptions are satisfied.

4 Simulations when the LL framework is correct

For data generation processes that satisfy the LL specification restrictions, the distribution of the LL trace test for cointegrating rank is as given in LL (and described briefly above). We were able to simulate the critical values implied by the distribution.[7] In particular, in these experiments we assume $N = 2$, a block-diagonal β matrix and the same cointegrating rank in all units of the panel. The homogeneity of the cointegrating vector is also assumed, for simplicity. Results for $N = 4$ and $N = 8$ are summarized in Section 6 below.

The data generation processes are as given in Table 3. DGP 1 is the simplest null, with rank zero in both units. DGP 2A has rank 1 in both units, with the loading matrix α constructed such that the equilibrium correction terms enters into only one equation of each unit. In DGP 2B, with rank 1 in both units, the first equilibrium correction term enters one equation each of *both* units whereas the second equilibrium correction term enters only one equation of the second unit. DGP 2A and DGP 2B are constructed in order to have the same amount of integration (as reflected in the absolute values of the non-zero roots of the companion matrix) while having different equilibrium correction properties. DGP 2C is a variant of DGP 2A with less integration.[8]

In Table 4 we report the size of not only the LL tests but also of the Johansen procedure unit by unit, jointly, and in full systems. Thus the column headed LL gives the rejection frequencies of the null hypothesis of $r_i = r_{LL} = 0$, $i = 1, 2$, at the 5% significance level when the LL statistic is calculated for DGP 1 and LL critical values are used. For DGP 2A to DGP 2C the rejection frequencies of the null hypothesis of $r_i = r_{LL} = 1$ at the 5% significance level is given, again with

[7] The values coincide with those in LL and, for the special case where the cointegrating rank is 0, with those presented in Johansen (1995).

[8] Other experiments with different configurations for α yielded qualitatively similar results, while with smaller stationary roots the statistics have better size and higher power.

Larsson and Lyhagen (1999) critical values.

The columns headed "unit 1" and "unit 2" provide the corresponding rejection frequencies, for each unit individually, at the 5% significance level using Johansen (1995) critical values. The rejection frequencies in the column headed "joint" are calculated by setting the significance level of the unit by unit tests at $1-(0.95)^{1/2}$, so that the significance of the joint test (i.e. the probably of rejecting rank 0 in both units in DGP 1, and rank 1 in both units in DGP 2A to 2C) is 5%. The column headed "J system " gives the rejection frequencies (of $r_J = 0$ for DGP 1 and $r_J = 2$ for DGPs 2A to 2C) when the full four-dimensional system, is estimated *without* restriction. Note that r_{LL} indicates the cointegrating rank in the LL sense, and r_J that in the full system.

The final column of Table 4 gives the rejection frequencies of cointegration among the common trends derived from the unit by unit cointegration analysis, whenever the null hypothesis is accepted in each unit. Since for all the data generation processes considered in Table 3 there is no cross-unit cointegration, the final column therefore reports the rejection frequencies of $r = 0$ whenever four common trends (two from each unit) are extracted in the samples generated by DGP 1, based on testing for cointegrating rank unit by unit, and whenever two common trends (one from each unit) are extracted in the samples generated by DGP 2A to 2C. This procedure is equivalent to that proposed by Gonzalo and Granger (1995) for checking for cointegration among sub-systems of cointegrated systems. Gonzalo and Granger suggested that the asymptotic distribution of the cointegration test is not affected by having estimated the common trends in the first stage. Hence, the critical values used are taken from Johansen (1995).

The results on size, as reported in Table 4, are encouraging. Except for the joint test, no substantial size distortions are evident. The distortions of the joint test disappear as the sample size increases to 400 but, particularly for DGP 2A and 2B, they are present and important at economically typical sample sizes of 100 or 200. The distortions are greater with more feedback (compare DGP 2B with 2A) and lower with lower integration (compare DGP 2A with 2C).

The Gonzalo and Granger procedure is slightly under-sized whenever cointegrating vectors are estimated, as in DGP 2A to 2C, even at large sample sizes of 400. This is likely a consequence of using critical values that are strictly not applicable with constructed as opposed to raw series. It is in principle possible to adjust the critical values for size but since such corrections would be *ad hoc*, the size distortions are not great, and the general points can be made in the absence of such size corrections, we do not pursue this further.

Table 5 (under identical headings) provides the power of the test procedures for DGP 2A to 2C. Thus for DGP 2A to 2C the rejection frequencies of $r_{LL} = 0$ are given in the column headed LL, while the remaining columns provide the same information for the other tests. All the tests have power approaching 1 as sample size increases. Broadly speaking however, the LL test is seen to have the best power properties for a majority of cases, both at low, medium and large sample sizes. The test based on estimating the Johansen full system is the least powerful (although increasing rapidly to 1) for all but one of the cases. This is to be expected, since when the LL null is satisfied, estimation of the Johansen full system involves estimating a large number of unnecessary parameters. This leads to a loss in efficiency in estimation of important parameters and loss of power.[9] Estimating the system unit by unit is partly beneficial (by cutting down on the number of parameters to be estimated) although cross-unit connections *via* the α matrix are not taken into account. The joint test therefore occupies the middle ground in terms of power performance.

Table 6 provides details of the bias and standard deviation of the estimates of

[9] See discussion of results of Table 6 below.

the cointegrating vectors for each of the three different methods. In terms of bias, even for T=100, the LL test and the Johansen trace statistic unit by unit perform quite well. The latter is slightly better than the former, also in terms of efficiency, *i.e.*, lower standard errors for the estimated coefficients. The full system Johansen ranks third for both bias and efficiency.

5 Simulations when the LL framework is not correct

From the previous section, the LL test appears to have the proper size, good power compared to other cointegration procedures, and to yield some efficiency gains when estimating the cointegrating coefficients. Yet, as we stressed in the introduction, the LL framework (and the even more restricted models in the panel unit root literature), may be inappropriate in applications with macroeconomic data. In particular, in this context there can exist cointegrating relationships across the units, and the units can be driven by a different number of common trends. We now evaluate the consequences of these two features on the performance of the LL test, and consider some "diagnostic tests" to spot their presence.

We use simulations to undertake this part of the analysis, instead of looking for analytical expressions of the distributions under the alternative. The difficulties of deriving densities in the presence of cross-unit cointegration are noted by Phillips (1999) who comment that when there are strong correlations among the units of a panel (as would apply in our cases of interest), standard central limit theory and laws of large numbers will no longer apply. Moreover our interest is not solely in the limit distributions (with N and T approaching infinity) but we wish also to look at the behaviour of these statistics for small N and medium-sized T. The predictions of asymptotic theory are thus of limited use here.

5.1 Cointegration across the units

When the units are related by cointegrating relationships, the hypothesis of a block diagonal β is violated. Several structures for β are now possible, and we focus on three of these in our simulation experiments. The data generation processes are described in Table 7. In DGP 3 there exists only one cointegrating relationship across the units, so that $r_{LL} = 0$ and $r_J = 1$. The DGPs 3A, 3B and 3C differ for the loading matrix, α, while in 3D there is more stationarity. In DGP 4 we consider the case $r_{LL} = 0$ and $r_J = 2$, namely, two cointegrating relationships across the units, and none within the units. In DGP 5 we also allow for within-unit cointegration, and keep one cross-unit relationship, so that $r_{LL} = 1$ and $r_J = 3$. Several sub-cases of DGP 4 and 5 are analyzed, which differ for the structure of the α matrix (indicated by A,B,C) and for the magnitude of the roots (D).

The rejection frequencies of the LL test are reported in the LL columns of Table 8. For DGP 3, where $r_{LL} = 0$, the probability of rejecting $r_{LL} = 0$ quickly increases towards one, while that of rejecting $r_{LL} = 1$ is in the range $0.26-0.28$ when $T = 400$. For DGP 4 also the probability of rejecting $r_{LL} = 1$ increases to one, and the same is true for DGP 5. These results indicate a massive rejection of the proper null hypothesis of the LL test in the presence of cross-unit cointegration. Cointegration across units is wrongly attributed to cointegration within units.

It is interesting to evaluate whether this is also a problem in the unit by unit analysis using the Johansen trace test. Indeed, from the columns “unit 1” and “unit 2” of Table 8, there are cases when the size of the test is severely biased upwards, even if the highest rejection probabilities of $r_{LL} = 0$ are much smaller than those from LL, about .38 for $T = 400$. The distortion is related to the cointegrating relationships affecting several variables of the system (cases B and C). As a consequence, the joint procedure based on the combination of the unit by unit results also presents size distortions in these cases (column "joint").[10]

In the light of these results, it seems particularly important either to use a full system approach, or at least to evaluate whether there is cross-unit cointegration. The former approach though is not feasible with a larger number of units, so that the latter becomes even more important. The last three columns of Table 8 report the size and power of the Gonzalo and Granger procedure. The power of the test (columns $r_C = 0$ for all DGPs and also $r_C = 1$ for DGP 4) is in general rather low for $T = 100$, but quickly increases with T and is often close to one for $T = 400$. The size (columns $r_C = 1$ for DGPs 3 and 5, and column $r_C = 2$ for DGP 4) is slightly lower than the nominal value also for $T = 400$, as in the previous section.

Overall, these results suggest that the LL test, and to a certain extent unit by unit analysis, can lead to over-acceptance of within-unit cointegration when there exists cross-unit cointegration. On the other hand, the latter is detected quite well by the Gonzalo and Granger test, at least for large enough sample sizes.

5.2 Different cointegrating ranks in each unit

The cases of different within-unit cointegrating ranks we consider are listed in Table 9. In DGP 6 it is $r_1 = 1$, $r_2 = 0$; for DGP 7 $r_1 = 2$, $r_2 = 0$; for DGP 8, $r_1 = 2$, $r_2 = 1$. For each DGP we also consider different loading matrices α (subcases A, B, C), and lower stationary roots (D).

The rejection frequencies of the LL test are reported in the LL columns of Table 10. When $T = 400$, $r_{LL} = 0$ is rejected with probability one in all cases, and also

[10]Note that, in order to reduce uncertainty, the same random numbers are used for all experiments. Hence some values for the unit by unit analysis are the same (e.g. those for unit 2, DGPs 3A, 3B, 3D, 4A, 4B, 4D).

$r_{LL} = 1$ is rejected with probability close to one for DGPs 7 and 8. When T is smaller, the rejection frequencies are lower in some cases, but overall these results also indicate a tendency of the LL test to substantially under reject within-unit cointegration.

The unit by unit cointegration analyses provide an indication about differences in the cointegrating ranks (columns "unit 1" and "unit 2" of Table 10). In fact when $T = 400$, the unit by unit tests in general have power close to one, and size close to the nominal level. The exceptions are DGPs 6C and 7C, when the size of the test is larger, about 0.14. This is the case where the cointegrating vectors from the first unit also affect the second unit.

When T is smaller, the power of the tests in some DGPs is low, in particular those of type A. In these cases the performance of the joint test (column "joint") is also unsatisfactory, with substantial over rejection of the correct null.

In summary, when the cointegrating ranks are different across units, the LL test tends to over-accept the presence of within-unit cointegration. In other words, cointegration in one unit biases the test towards non rejection of cointegration in the other unit. Unit by unit analysis is accurate when T is large enough, and can provide a useful diagnostic test for different cointegrating ranks across units.

The suggestion for empirical analysis that emerges from this section is to use full-system estimation whenever possible. If this is not feasible, the first step should be a unit by unit cointegration analysis. The second step is to test for the absence of cross-unit cointegration by means of the Gonzalo and Granger procedure. If this is accepted, and the unit by unit analysis does not indicate the presence of different ranks across units, the third step is to apply the LL statistic, that can yield efficiency gains in terms of higher power and lower standard errors for the estimated cointegrating coefficients. This sequential empirical procedure is implemented in

Section 7, to evaluate in more detail the presence of cointegration in the interest rate data.

6 Summary of results for more units

Since the case of $N = 2$ might be thought to be unduly restrictive, in this section we summarize the results of simulations when the number of units in the panel is increased to 4 or 8. Conventional panel studies often consider cases where the number of units is even larger, and a limitation of the highly parameterized maximum likelihood methods considered in this paper is that this cannot be done, unless T is very large with respect to N or the dependence across the units is severely restricted. Imposing these restrictions leads to the framework considered by Larsson et al. (1998), or by Kao (1999) and Pedroni (1999) inter alia, where the latter papers make the further assumption of one cointegrating vector. The tradeoff between higher dimensionality and a priori restrictions is an issue that merits further investigation. Nevertheless, a panel of 4 or 8 units is a reasonable size when considering for example data from the G7 countries, or from the largest economies of the European Union, or economic groupings such as NAFTA. The key economic indicators of the main sectors of an economy, of geographically differentiated regions within a country, or of the different segments of the labour market could also be investigated within this context.

In order to evaluate the size and power of the competing approaches, we made use of DGPs which replicated the structures used for $N = 2$. Thus, for example, DGP 2A for $N = 4$, took the form

$$\alpha = \begin{pmatrix} -.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -.1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -.1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -.1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The extensions for the other DGPs studies, namely, 3 to 8, are similarly straightforward. Moreover, for DGPs 3-5 we also considered cases with more than two cointegrating vectors across the units, and for DGPs 6-8 different degrees of cointegration whithin each unit. A detailed account of these results (available upon request) is not presented here for space constraints. The following important features are however worth noting:

1) As the number of units in increased, the size of the LL tests becomes distorted if asymptotic critical values are used. For example, when data are generated with $r_i = 1$ for each i, $N = 4$, $T = 200$, and the largest stationary roots are 0.8, the LL test rejects the null with frequency of 0.19 at a confidence level of 5%. This distortion becomes very severe for $N = 8$, with the rejection frequency increasing to 0.90. The distortion also increases with the magnitude of the largest stationary roots, if the other features of the DGP are left unchanged. Thus, with $r_i = 1$ for each i, $N = 4$, $T = 200$, and largest stationary roots of 0.9, the corresponding rejection frequency is 0.32.

The issue of size distortions is evident in the discussion in LL of their simulations, leading them to suggest the use of "Bartlett corrections" for the test statistic. The form of the correction is given by

$$\widetilde{C}_T = \frac{E(C_T)}{E(C_\infty)} C_\infty,$$

where C_∞ is the asymptotic critical value and $E(C_T)$ and $E(C_\infty)$ are, respectively, the expectations of the finite sample and asymptotic distributions of the test statistic. Both $E(C_T)$ and $E(C_\infty)$ can be approximated by simulations. Our results suggest that the use of $\widetilde{C}_T$ is effective in correcting the size of the tests. In other words, $\widetilde{C}_T$ is close in magnitude to the corresponding empirical quantile of the distribution of the test statistic. The power of the test when using $\widetilde{C}_T$ is satisfactory, close to one even for T=200.

2) We also investigated the performance of the LL procedure in the presence of cross-unit cointegration and different cointegrating ranks. As before, we find that even the presence of few cross-unit cointegrating relationships substantially biases the LL test towards rejection of no within-unit cointegration, e.g., with $T = 200$, $r_i = 0$ for each i, and one cointegrating relationship among all units, the probability of rejecting the null $r_i = 0$ for each i is 0.56 when $N = 8$. This can be compared with the figure of 0.62 for DGP 3A in Table 8. We conclude from this that the importance of cross-unit cointegrating relationships does not decrease even when the number of units increases. For the case of different ranks across the units, the over-acceptance of within-unit cointegration reported in Table 10 and discussed in Section 5.2 is confirmed. The degree of over-acceptance increases with the number of units which have cointegration. For example, if $N = 4$, $T = 200$, the rejection frequencies of $r_i = 0$ for each i are 0.18, 0.44, and 0.69 for, respectively, $(r_1 = 1, r_2 = r_3 = r_4 = 0)$, $(r_1 = r_2 = 1, r_3 = r_4 = 0)$, and $(r_1 = r_2 = r_3 = 1, r_4 = 0)$.

3) As far as the various versions of the Johansen test are concerned (unit by unit, joint, and system), these retain their good size and power properties when the units do not cointegrate with each other. In the presence of cross-unit cointegration, however, as for $N = 2$, the unit by unit analysis (and therefore the joint test) can reject the null of no cointegration too often. The size of the GG test remains slightly lower than the nominal value, at about 0.04. Its power is relatively low when there is one cross-unit relationship, but quickly increases with the amount of cross-unit cointegration and the number of observations.

In summary, it is satisfying to conclude that once the Bartlett corrected critical values are used for the LL test, the overall performance of the various tests considered remains qualitatively unaltered from that reported in detail for $N = 2$.

7 Empirical Examples Revisited

The first step in the sequential procedure we suggested above is unit by unit analysis. Hence, we specify similar models for each country to estimate the unit-specific cointegrating vectors. We then complement the analysis by extracting the common trends from each unit and looking for cointegration among these. Finally, we run the LL test to evaluate whether the outcome is in line with the simulation results. Detailed results are reported for $N = 2$, and a summary of the main findings is presented for $N = 8$.

7.1 Germany and The Netherlands

The chosen specification for Germany involves 4 lags and an unrestricted constant. The available sample period goes from May 1990 to January 1999, for a total sample size of 105 observations. Diagnostics tests for this specification point towards its congruence since only heteroskedasticity and normality of the residuals of the short-term rate are marginally rejected. Cointegration tests for each rank are reported in Table 11, first panel, along with the corresponding critical values.

The statistics indicate that the cointegrating rank in the model for Germany is zero. One can then use both the short-and the long-term rate for Germany in a successive step of the analysis based on the Gonzalo and Granger method.

The chosen specification for the Netherlands involves 4 lags and an unrestricted constant. The available sample period goes from May 1990 to January 1999, for a total sample size of 105 observations. Diagnostics tests for this specification indicate no evidence of misspecification. Cointegration tests for each rank are reported in Table 11, second panel, along with the corresponding critical values.

The statistics indicate that the cointegrating rank in the model for The Netherlands is also zero, which is consistent with the full-system cointegration analysis. As a consequence, the analysis for cointegration among German and Dutch interest rates would have to be conducted on the full system. From the results in Section 3, we know that there exists one cross country cointegrating relationship. Hence, we expect the LL test to be biased toward rejection of the null hypothesis $r_{LL} = 0$.

For the LL model the sample goes from May 1990 to January 1999, we leave the constant unrestricted, and use 4 lags. The panel LR tests are reported in Table 12, first panel.

In line with our expectations, the table indicates rejection of $r_{LL} = 0$ at the 10%, though there is marginal non-rejection of the hypothesis at the 5%. Notice also that whatever choice the investigator makes on the basis of the LL test, the resulting model will be misspecified. If $r_{LL} = 0$ is chosen, the system will be analyzed in first differences, thereby overlooking the existence of cointegration between the two countries. If instead $r_{LL} = 1$ is selected, the wrong cointegrating relationships will be included in the ECM models.

7.2 Germany and Denmark

The first tentative specification for Denmark involves 4 lags and an unrestricted constant. The available sample period goes from May 1990 to January 1999, for a total sample size of 105 observations. We are led to perform the final analysis on the reduced sample starting in October 1993 because large outliers in 1993 provoke heteroskedasticity and non-normality in the Danish short-term rate. There is no evidence of misspecification in the reduced sample going from October 1993 to January 1999, except for non-normality in the short-term rate.

The cointegration properties do not change in the two samples, since the rank in both is found to be zero. Both sets of results are reported in Table 11, third and fourth panels. The finding of zero cointegrating rank in Denmark implies that the Gonzalo and Granger analysis for cointegration among German and Danish interest rates would also have to be conducted on the full system. From Section 3, we know that there exists only one cross country cointegrating relationship, so that the LL test should again over-reject the null of no cointegration.

To avoid the troublesome period, we only estimate the LL model on the second subsample, with 4 lags and an unrestricted constant. The panel LR tests are reported in Table 12, second panel. This time the LL test rejects a common unit-by unit rank of zero at 5%, as well as a common rank of one. This would lead the investigator to estimate the model imposing a common rank of 2, with no cross-country cointegration. Since the only cointegration vector in the full system involves variables from both countries, the results would probably be very misleading.

7.3 More units

We now increase the number of units in the panel to 8, by including France, Italy, UK, Spain and Austria in the analysis. As above, we start by analyzing the full system, which includes the short- and long-term interest rates for each country, giving a total of 16 variables. Only one lag was included in the unrestricted VAR, both for the sake of parsimony and because the resulting model did not present any major evidence of misspecification.[11] The Johansen trace test suggests the presence of 9 cointegrating relationships across the variables. We do not attempt to identify them, but we note that this outcome already suggests that either some units have cointegrating rank larger than one or that there exist cross-unit cointegrating relationships.

In order to evaluate the first possibility, we performed a unit-by-unit cointegration analysis. In line with our results for Germany, The Netherlands and Denmark, there appears to be no within-unit cointegration in the other countries, with the possible exception of Spain. Thus the first requirement for the application of the LL test is satisfied. However, we still need to verify the presence or absence of cross-unit cointegration.

Given that the null hypothesis of zero rank in each unit is accepted, the GG test can be run with the original set of variables and is equivalent to the trace test in the full-system analysis. The outcome of such a test can now be interpreted as indicating the presence of many cross-unit cointegrating relationships. From the simulation experiments in Section 5.1 and 6, we know that the presence of cross-unit cointegration can substantially bias the LL test towards rejection of the no cointegration hypothesis ($r_i = 0$ for each i, which is supported by the unit-by-unit analysis) even when the number of units is large. The bias is in fact more serious than that caused by the violation of non-uniform rank in each individual unit.

Using corrected critical values calibrated to our model ($N = 8$, $T = 100$, and the largest stationary roots are 0.6), we find that the hypothesis $r_i = 0$ for each i is rejected at the 10% level, and $r_i = 1$ for each i is also strongly rejected. Thus,

[11]Full details of the estimations reported in this section are available from us upon request.

we have another example where the application of the panel cointegration tests is problematic because the main underlying hypothesis of no long run dependence across the units is violated, leading us to "spuriously" conclude in favour of the presence of within-unit cointegration.

8 Conclusions

Using data for short- and long-term interest rates in a small panel of OECD countries, the restrictions implied by the block diagonality of β that characterize the panel cointegration framework (not only in systems methods such as LL but also in the panel analogue of Engle-Granger methods as developed by Kao (1999) and Pedroni (1999) *inter alia)* can be very easily shown to be violated. Since we suspect that this is very likely to be the case when analyzing macroeconomic time series across countries, we should be careful in using these methods.

Our simulation results for $N = 2, 4$ and 8 indicate that when the hypotheses underlying the LL framework are satisfied, the LL test has good size and power properties, and often yields gains in efficiency relative to full system analysis for estimation of the cointegrating parameters.

The consequences of violations of the LL assumptions can however be serious. Our results suggest that the LL test can lead to substantial over-acceptance of within-unit cointegration when there exists cross-unit cointegration or when the ranks are different across units. Since tests for unit roots in panels and other existing tests for cointegration can be re-interpreted within the generalized system framework, it follows that such tests also reject non-stationarity too often. Thus, the common empirical finding in the PPP and convergence literature, when tested in panels of countries, namely that PPP holds and country GDPs converge, may be a spurious consequence of such violations. We are taking a closer look at this issue in on-going research.

Unit by unit analysis however is accurate when T is large enough, and can provide a useful diagnostic test for different cointegrating ranks across units, while cross-unit cointegration is detected by the Gonzalo and Granger test, at least for large enough sample sizes.

The suggestion for empirical analysis that therefore emerges is to use full system estimation whenever possible. If this is not feasible, the first step should be a unit by unit cointegration analysis. The second step is to test for the presence of cross unit cointegration by means of the Gonzalo and Granger procedure. If this is rejected, and the unit by unit analysis does not indicate the presence of different ranks across units, the third step is to apply the LL statistic since this can yield efficiency gains in terms of higher power and lower standard errors for the estimated cointegrating coefficients.

References

Doornik, J. A. (1999): *Object-Oriented Matrix Programming Using Ox, 3rd Ed.* Timberlake Consultants Press, London. www.nuff.ox.ac.uk/Users/Doornik.

Doornik, J. A., and D. F. Hendry (1997): *Modelling Dynamic Systems Using PcFiml 9 for Windows.* Timberlake Consultants, London.

Engle, R. F.and Granger, C. W. (1987): "Cointegration and Error-Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

Evans, P., and G. Karras (1996): "Convergence Revisited," *Journal of Monetary Economics*, 37, 249–265.

Frankel, J., and A. Rose (1996): "A Panel Project on Purchasing Power Parity: Mean Reversion Between and Within Countries," *Journal of International Economics*, 40, 209–224.

Gonzalo, J. (1994): "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," *Journal of Econometrics*, 60, 203–233.

Gonzalo, J., and C. W. J. Granger (1995): "Estimation of Common Long-Memory Components in Cointegrated Systems," *Journal of Business and Economic Statistics*, 13, 27–36.

Groen, J. J., and F. R. Kleibergen (1999): "Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models," Discussion Paper Discussion Paper 55, Tinbergen Institute.

Im, K. S., P., and Y. M. H.and Shin (1997): "Testing for Unit Roots in Heterogeneous Panels," Discussion paper, Department of Applied Economics, University of Cambridge.

JOHANSEN, S. (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford and New York.

KAO, C. (1999): "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics*, 90, 1–44.

LARSSON, R., AND J. LYHAGEN (1999): "Likelihood-Based Inference in Multivariate Panel Cointegration Models," Discussion Paper Working Paper Series in Economics and Finance 331, Stockholm School of Economics and Finance.

——— (2000): "Testing for Common Cointegrating Rank in Dynamic Panels," Discussion Paper Working Paper Series in Economics and Finance 378, Stockholm School of Economics.

LARSSON, R., J. LYHAGEN, AND M. LÖTHGREN (1998): "Likelihood-Based Cointegration Tests in Heterogeneous Panels," Discussion Paper Working Paper Series in Economics and Finance 250, Stockholm School of Economics.

LEE, K. E., M. H. PESARAN, AND R. SMITH (1997): "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model," *Journal of Applied Econometrics*, 12, 357–392.

LEVIN, A., AND C. F. LIN (1993): "Unit Root Tests in Panel Data: New Results," Discussion Paper Discussion Paper 56, University of California at San Diego.

MCKOSKEY, S., AND C. KAO (1998): "A Residual-Based Test for the Null of Cointegration in Panel Data," *Econometric Reviews*, 17, 57–84.

OCONNELL, P. (1998): "The overvaluation of purchasing power parity," *Journal of International Economics*, 44, 1–19.

PAPELL, D. (1997): "Searching for Stationarity: Purchasing Power Parity under the Current Float," *Journal of International Economics*, 43, 313–323.

Pedroni, P. (1999): "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors," *Oxford Bulletin of Economics and Statistics*, 61 Special Issue, 653–70.

Phillips, P. C. B.and Moon, H. R. (1999): "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057–1111.

Appendix

Table 1: Cointegration analysis for Germany and the Netherlands

Ho: rank = r	max eigenvalue	95%	trace	95%
$r = 0$	33.16**	27.1	46.74	47.2
$r \leq 1$	10.45	21.0	13.58	29.7
$r \leq 2$	3.117	14.1	3.127	15.4
$r \leq 3$	0.0107	3.8	0.0107	3.8

Variables	α **coefficients** (**standard errors**)	β **coefficients**
Germany short	-0.28589 (0.052418)	1
Germany long	0	0
Netherlands short	0	-1
Netherlands long	0	0
LR test of restrictions: $\chi^2_{(6)} = 4.8436$ [0.5640]		

Table 2: Cointegration analysis for Germany and Denmark

Ho: rank=r	Max eigenvalue	95%	Trace	95%
$r = 0$	40.04**	27.1	56.77**	47.2
$r \leq 1$	19.65	21.0	22.83	29.7
$r \leq 2$	7.209	14.1	7.288	15.4
$r \leq 3$	0.07885	3.8	0.07885	3.8

Variables	α **coefficients** (**standard errors**)	β **coefficients**
Germany short	-0.043439 (0.018386)	-1.19
Germany long	0	0
Denmark short	-0.48171 (0.11189)	1
Denmark long	-0.091265 (0.027468)	0
LR test of restrictions: $\chi^2_{(3)} = 4.9176$ [0.1779]		

Table 3: Data Generating Processes for Size and Power: Block-diagonal β

DGP	r_{LL}	r_J	α	β'	*Roots*
1	0	0	$\mathbf{0}$	$\mathbf{0}$	(1,1,1,1)
2A	1	2	$\begin{pmatrix} -0.1 & 0 \\ 0 & 0 \\ 0 & -0.1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$	(1,1,0.9,0.9)
2B	1	2	$\begin{pmatrix} -0.1 & 0 \\ 0 & 0 \\ 0.1 & -0.1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$	(1,1,0.9,0.9)
2C	1	2	$\begin{pmatrix} -0.2 & 0 \\ 0 & 0 \\ 0 & -0.2 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$	(1,1,0.8,0.8)

r_{LL}: Cointegrating rank for each unit in sense of Larsson and Lyhagen

r_J: Cointegrating rank for full system in sense of Johansen

Table 4: Size of tests with block diagonal β matrix

DGP	T	LL$^{(a)}$	J unit by unit$^{(b)}$			J system $^{(d)}$	GG $^{(e)}$
			unit 1	unit 2	joint$^{(c)}$		
1	100	0.064	0.048	0.058	0.056	0.065	0.051
	200	0.059	0.051	0.051	0.056	0.060	0.045
	400	0.057	0.052	0.051	0.052	0.057	0.045
2A	100	0.098	0.048	0.049	0.815	0.022	0.021
	200	0.087	0.055	0.052	0.142	0.051	0.024
	400	0.069	0.051	0.052	0.051	0.056	0.028
2B	100	0.118	0.049	0.035	0.887	0.024	0.038
	200	0.093	0.055	0.046	0.384	0.045	0.033
	400	0.069	0.051	0.048	0.055	0.055	0.031
2C	100	0.098	0.054	0.054	0.136	0.057	0.027
	200	0.072	0.054	0.050	0.054	0.058	0.029
	400	0.060	0.050	0.053	0.050	0.053	0.031

(a): Rejection frequencies of null hypothesis at 5% significance level using LL(1999) asymptotic critical values, when LL statistic is calculated.

(b): Rejection frequencies, unit by unit, of null hypothesis at 5% significance level, using Johansen (1995) asymptotic critical values.

(c): Calculated by setting the significance level of the unit by unit Johansen tests at $1 - (0.95)^{1/2}$, so that the significance of the joint test is 5%.

(d): Rejection frequencies at 5% significance level using Johansen (1995) asymptotic critical values when system is estimated without restriction (apart from those necessary for identification).

(e): Rejection frequencies of cointegration among the common trends derived from the unit by unit cointegration analysis when the null hypothesis is accepted in each unit.

Table 5: Power of tests with block diagonal β matrix

DGP	T	LL[(a)]	J unit by unit[(b)]			J system [(d)]
			unit 1	unit 2	joint[(c)]	
2A	100	0.513	0.576	0.575	0.229	0.136
	200	0.978	0.982	0.985	0.914	0.635
	400	1.000	1.000	1.000	1.000	0.999
2B	100	0.699	0.576	0.367	0.151	0.165
	200	0.999	0.982	0.792	0.667	0.522
	400	1.000	1.000	0.998	0.995	0.974
2C	100	0.983	0.985	0.983	0.917	0.658
	200	1.000	1.000	1.000	1.000	0.999
	400	1.000	1.000	1.000	1.000	1.000

$(a) - (d)$: See notes to Table 4

Table 6: Estimates of β with block diagonal β matrix

DGP	T		LL[(a)]		J unit by unit[(b)]		J system [(c)]	
			$\widehat{\frac{\beta_{12}}{\beta_{11}}}$	$\widehat{\frac{\beta_{22}}{\beta_{21}}}$	$\widehat{\frac{\beta_{12}}{\beta_{11}}}$	$\widehat{\frac{\beta_{22}}{\beta_{21}}}$	$\widehat{\frac{\beta_{12}}{\beta_{11}}}$	$\widehat{\frac{\beta_{22}}{\beta_{21}}}$
2A	100	mean	1.169	0.961	1.058	1.001	1.141	1.005
		s.d.	8.837	10.730	2.716	2.783	17.150	20.564
	200	mean	1.014	1.029	1.002	1.004	1.011	0.871
		s.d.	1.026	3.020	0.159	0.161	4.779	5.605
	400	mean	0.999	1.000	0.999	1.000	0.781	1.052
		s.d.	0.068	0.070	0.065	0.064	20.829	18.631
2B	100	mean	1.052	1.058	1.058	1.371	1.092	0.937
		s.d.	2.884	7.068	2.716	33.670	9.922	18.448
	200	mean	1.002	0.823	1.002	1.036	1.150	0.540
		s.d.	0.126	18.062	0.160	4.810	10.389	15.184
	400	mean	1.000	1.000	0.999	1.045	1.007	0.731
		s.d.	0.052	0.075	0.065	0.126	1.117	1.437
2C	100	mean	1.004	1.015	1.002	1.005	0.772	0.818
		s.d.	0.334	0.839	0.153	0.173	23.713	6.589
	200	mean	1.000	0.999	1.000	1.000	0.942	0.912
		s.d.	0.069	0.112	0.066	0.063	3.276	3.499
	400	mean	1.000	1.000	1.000	1.000	1.363	0.496
		s.d.	0.032	0.030	0.031	0.030	38.559	37.520

(a): Monte Carlo mean and standard deviation of estimated normalised cointegrating parameter in each unit using LL method.

(b): Monte Carlo mean and standard deviation of estimated normalised cointegrating parameter in each unit using Johansen unit by unit.

(c): Monte Carlo mean and standard deviation of estimated normalised cointegrating parameter in each unit using Johansen full system.

Table 7: Description of Data Generating Processes for Rejection Frequencies: Non-block-diagonal β

DGP	r_{LL}	r_J	α'	β'	*Roots*
3A	0	1	$\begin{pmatrix} -0.1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,1,0.9)
3B	0	1	$\begin{pmatrix} -0.1 & -0.1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,1,0.9)
3C	0	1	$\begin{pmatrix} -0.1 & 0 & 0 & -0.1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,1,0.9)
3D	0	1	$\begin{pmatrix} -0.2 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,1,0.8)
4A	0	2	$\begin{pmatrix} 0 & -0.1 & 0 & 0 \\ -0.1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,0.9,0.9)
4B	0	2	$\begin{pmatrix} 0 & -0.1 & 0 & 0 \\ -0.1 & -0.1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,0.9,0.9)
4C	0	2	$\begin{pmatrix} 0 & -0.1 & 0 & 0 \\ -0.1 & 0 & -0.1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,0.9,0.9)
4D	0	2	$\begin{pmatrix} 0 & -0.2 & 0 & 0 \\ -0.2 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}$	(1,1,0.8,0.8)
5A	1	3	$\begin{pmatrix} -0.1 & 0 & 0 & 0 \\ 0 & 0 & -0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$	(1,0.9,0.9,0.9)
5B	1	3	$\begin{pmatrix} -0.1 & 0 & 0 & -0.1 \\ 0 & 0 & -0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$	(1,0.9,0.9,0.9)
5C	1	3	$\begin{pmatrix} -0.1 & 0 & 0 & -0.1 \\ 0 & 0 & -0.1 & 0 \\ 0 & 0 & -0.1 & 0.1 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$	(1,0.9,0.9,0.9)
5D	1	3	$\begin{pmatrix} -0.2 & 0 & 0 & 0 \\ 0 & 0 & -0.2 & 0 \\ 0 & 0 & 0 & 0.2 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$	(1,0.8,0.8,0.8)

Table 8: Rejection frequencies of tests with non-block diagonal β matrix

DGP	T	LL[(a)]		J unit by unit[(b)]					GG [(d)]		
				unit 1		unit 2		joint			
		r_{LL}=0	r_{LL}=1	r_J=0	r_J=1	r_J=0	r_J=1		r=0	r=1	r=2
3A	100	0.231	0.105	0.054	0.004	0.054	0.006	0.055	0.210	0.020	-
	200	0.621	0.235	0.051	0.004	0.053	0.005	0.056	0.605	0.036	-
	400	0.993	0.281	0.053	0.003	0.051	0.004	0.050	0.993	0.041	-
3B	100	0.583	0.221	0.263	0.026	0.054	0.006	0.219	0.502	0.037	-
	200	0.970	0.288	0.324	0.026	0.053	0.005	0.276	0.960	0.039	-
	400	1.000	0.275	0.369	0.028	0.051	0.005	0.320	1.000	0.038	-
3C	100	0.580	0.218	0.054	0.004	0.265	0.023	0.221	0.500	0.035	-
	200	0.969	0.285	0.051	0.004	0.341	0.030	0.288	0.960	0.042	-
	400	1.000	0.264	0.053	0.003	0.382	0.029	0.333	1.000	0.031	-
3D	100	0.628	0.236	0.056	0.005	0.054	0.006	0.067	0.613	0.038	-
	200	0.995	0.285	0.060	0.005	0.053	0.005	0.059	0.994	0.040	-
	400	1.000	0.271	0.059	0.003	0.051	0.005	0.052	1.000	0.038	-
4A	100	0.512	0.350	0.063	0.008	0.054	0.006	0.058	0.495	0.115	0.016
	200	0.977	0.937	0.063	0.006	0.053	0.005	0.058	0.976	0.632	0.038
	400	1.000	1.000	0.058	0.006	0.051	0.005	0.045	1.000	1.000	0.033
4B	100	0.696	0.416	0.164	0.025	0.054	0.006	0.126	0.663	0.121	0.014
	200	0.998	0.861	0.222	0.034	0.053	0.005	0.171	0.997	0.470	0.026
	400	1.000	1.000	0.264	0.034	0.051	0.005	0.197	1.000	0.966	0.020
4C	100	0.843	0.638	0.142	0.023	0.265	0.023	0.241	0.798	0.240	0.020
	200	0.997	0.993	0.180	0.026	0.341	0.302	0.306	0.999	0.837	0.026
	400	1.000	1.000	0.203	0.026	0.382	0.029	0.358	1.000	1.000	0.021
4D	100	0.979	0.939	0.066	0.006	0.054	0.006	0.057	0.978	0.630	0.036
	200	1.000	1.000	0.070	0.005	0.053	0.005	0.054	1.000	0.999	0.032
	400	1.000	1.000	0.070	0.007	0.051	0.005	0.047	1.000	1.000	0.026

(continued)

5A	100	0.771	0.364	0.584	0.049	0.595	0.061	0.803	0.520	0.039	-
	200	1.000	0.842	0.984	0.056	0.986	0.062	0.131	0.956	0.023	-
	400	1.000	1.000	1.000	0.051	1.000	0.058	0.041	1.000	0.018	-
5B	100	0.954	0.424	0.584	0.049	0.715	0.068	0.751	0.714	0.024	-
	200	1.000	0.896	0.984	0.056	0.993	0.083	0.121	0.905	0.013	-
	400	1.000	1.000	1.000	0.051	1.000	0.103	0.068	1.000	0.009	-
5C	100	0.993	0.687	0.584	0.049	0.911	0.172	0.681	0.698	0.016	-
	200	1.000	0.997	0.984	0.056	0.999	0.221	0.204	0.928	0.006	-
	400	1.000	1.000	1.000	0.051	1.000	0.270	0.199	1.000	0.003	-
5D	100	1.000	0.849	0.985	0.053	0.991	0.071	0.120	0.956	0.023	-
	200	1.000	1.000	1.000	0.054	1.000	0.063	0.046	1.000	0.020	-
	400	1.000	1.000	1.000	0.049	1.000	0.061	0.039	1.000	0.017	-

(a): Rejection frequencies of $r_{LL} = 0$ or 1 at 5% significance level using LL(1999) asymptotic critical values, when LL statistic is calculated

(b): Rejection frequencies, unit by unit, of rank 0 or 1 at 5% significance level, using Johansen (1995) asymptotic critical values.

(c): Rejection frequencies of null hypothesis using joint test, where the significance level of the unit by unit Johansen tests is set at $1 - (0.95)^{1/2}$.

(d) : Rejection frequencies of cointegration among common trends derived from unit by unit analysis when the null hypothesis is accepted in each unit.

Table 9: Description of Data Generating Processes for Rejection Frequencies: Block diagonal β matrix but different rank in each unit

DGP	r_J	α'	β'	*Roots*
6A	1	$\begin{pmatrix} -0.1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix}$	(1,1,1,0.9)
6B	1	$\begin{pmatrix} -0.2 & -0.1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix}$	(1,1,1,0.9)
6C	1	$\begin{pmatrix} -0.1 & 0 & -0.1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix}$	(1,1,1,0.9)
6D	1	$\begin{pmatrix} -0.2 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix}$	(1,1,1,0.8)
7A	2	$\begin{pmatrix} -0.1 & 0 & 0 & 0 \\ 0 & -0.1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$	(1,1,0.9,0.9)
7B	2	$\begin{pmatrix} -0.1 & 0.1 & 0 & 0 \\ 0 & -0.1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$	(1,1,0.9,0.9)
7C	2	$\begin{pmatrix} -0.1 & 0.1 & 0 & 0 \\ 0 & -0.1 & 0.1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$	(1,1,0.9,0.9)
7D	2	$\begin{pmatrix} -0.2 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$	(1,1,0.8,0.8)
8A	3	$\begin{pmatrix} -0.1 & 0 & 0 & 0 \\ 0 & -0.1 & 0 & 0 \\ 0 & 0 & -0.1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	(1,0.9,0.9,0.9)
8B	3	$\begin{pmatrix} -0.1 & 0.1 & 0 & 0 \\ 0 & -0.1 & 0 & 0 \\ 0 & 0 & -0.1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	(1,0.9,0.9,0.9)
8C	3	$\begin{pmatrix} -0.1 & 0.1 & 0 & 0 \\ 0 & -0.1 & 0.1 & 0 \\ 0 & 0 & -0.1 & 0.1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	(1,0.9,0.9,0.9)
8D	3	$\begin{pmatrix} -0.2 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 \\ 0 & 0 & -0.2 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	(1,0.8,0.8,0.8)

Table 10: Rejection frequencies of tests with block diagonal β matrix but different rank in each unit

DGP	T	LL[(a)]		J unit by unit[(b)]				
				unit 1		unit 2		joint[(c)]
		r_{LL}=0	r_{LL}=1	r_J=0	r_J=1	r_J=0	r_J=1	
6A	100	0.224	0.062	0.584	0.049	0.443	0.022	0.592
	200	0.615	0.093	0.984	0.056	0.053	0.005	0.101
	400	0.994	0.096	1.000	0.051	0.051	0.004	0.051
6B	100	0.962	0.126	1.000	0.055	0.058	0.006	0.060
	200	1.000	0.108	1.000	0.054	0.053	0.005	0.059
	400	1.000	0.099	1.000	0.055	0.051	0.005	0.054
6C	100	0.587	0.115	0.584	0.049	0.131	0.017	0.618
	200	0.969	0.116	0.984	0.056	0.138	0.015	0.157
	400	1.000	0.102	1.000	0.051	0.145	0.014	0.127
6D	100	0.635	0.100	0.985	0.053	0.054	0.006	0.096
	200	0.995	0.094	1.000	0.054	0.053	0.005	0.058
	400	1.000	0.090	1.000	0.049	0.051	0.005	0.050
7A	100	0.131	0.056	0.603	0.407	0.054	0.006	0.788
	200	0.499	0.306	0.998	0.987	0.053	0.005	0.094
	400	0.997	0.981	1.000	1.000	0.051	0.005	0.027
7B	100	0.218	0.079	0.796	0.417	0.054	0.006	0.767
	200	0.744	0.328	1.000	0.922	0.053	0.005	0.228
	400	1.000	0.866	1.000	1.000	0.051	0.005	0.026

(continued)

7C	100	0.493	0.232	0.796	0.417	0.119	0.014	0.769
	200	0.972	0.814	1.000	0.922	0.137	0.015	0.259
	400	1.000	1.000	1.000	1.000	0.143	0.014	0.102
7D	100	0.518	0.322	0.997	0.985	0.054	0.006	0.089
	200	0.997	0.984	1.000	1.000	0.053	0.005	0.031
	400	1.000	1.000	1.000	1.000	0.051	0.005	0.027
8A	100	0.225	0.086	0.603	0.407	0.209	0.031	0.975
	200	0.849	0.410	0.998	0.987	0.657	0.048	0.559
	400	1.000	0.982	1.000	1.000	0.999	0.053	0.038
8B	100	0.334	0.114	0.796	0.417	0.209	0.031	0.970
	200	0.949	0.365	1.000	0.922	0.657	0.048	0.625
	400	1.000	0.865	1.000	1.000	0.999	0.053	0.037
8C	100	0.475	0.227	0.796	0.417	0.094	0.017	0.985
	200	0.989	0.810	1.000	0.922	0.235	0.033	0.892
	400	1.000	1.000	1.000	1.000	0.737	0.051	0.452
8D	100	0.864	0.422	0.997	0.985	0.670	0.052	0.545
	200	1.000	0.984	1.000	1.000	0.999	0.051	0.033
	400	1.000	1.000	1.000	1.000	1.000	0.052	0.027

$(a)-(c)$: See notes to Table 8

Table 11: Unit by unit cointegration analysis

Ho: rank=r	Max eigenvalue	95%	Trace	95%
Germany, 1990-1999				
$r = 0$	4.86	14.1	4.886	15.4
$r \leq 1$	0.02539	3.8	0.025	3.8
Ho:rank=r	Max eigenvalue	95%	Trace	95%
The Netherlands, 1990-1999				
$r = 0$	2.984	14.1	3.087	15.4
$r \leq 1$	0.1033	3.8	0.095	3.8
Ho: rank=r	Max.eigenvalue	95%	Trace	95%
Denmark, 1990-1999				
$r == 0$	3.306	14.1	3.967	15.4
$r \leq 1$	0.2798	3.8	0.309	3.8
Ho: rank=r	Max.eigenvalue	95%	Trace	95%
Denmark, 1993-1999				
$r = 0$	7.596	14.1	8.175	15.4
$r \leq 1$	0.5791	3.8	0.579	3.8

Table 12: Panel LL tests

rank	**LR test**	**95% critical values**
Germany - The Netherlands, 1990-1999		
$r_{LL} = 0$	46.737	47.49
$r_{LL} = 1$	11.658	18.21

rank	**LR test**	**95% critical values**
Germany - Denmark, 1993-1999		
$r_{LL} = 0$	62.748	47.49
$r_{LL} = 1$	22.888	18.21