Testing Nonstationary Panels under Cross-Correlation:

New Methods and Empirical Evidence

Inaugural-Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Wirtschaftswissenschaften
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

vorgelegt von
Dipl.-Math. Adina Ioana Tarcolea
aus Bukarest

2006

Erstgutacher:	Prof. Dr. Uwe Hassler
Zweitgutachter:	Prof. Dr. Michael Binder
Tag der Promotion:	06.11.2006

List of original working papers

This thesis consists of the following original working papers.

1. Testing for Linear Trends in Dependent Heterogeneous Panels

Submitted.

2. Panel Cointegration Testing using Nonlinear Instruments

Co-authored by dr. Matei Demetrescu, submitted.

3. Combining Significance of Correlated Statistics with Application to Panel Data

Co-authored by dr. Matei Demetrescu and prof. dr. Uwe Hassler, *Oxford Bulletin of Economics and Statistics* **68** (2006), 647-633.

4. Combining Multi-Country Evidence on Unit Roots: The Case of Long-Term Interest Rates

Co-authored by prof. dr. Uwe Hassler, *Applied Economics Quarterly* **51** (2005), 181-189.

Abstract

This thesis is concerned with the derivation of new methods for the analysis of nonstationary, cross-correlated panels. The suggested procedures are carefully quantified by means of Monte Carlo experiments. Typical applications of the developed methods consist in multi-country studies, with several countries observed over a couple of decades. The empirical applications implemented here are the testing for trends in the investment share in European GDPs and the examination of OECD interest rates.

In the first chapter, a panel test for the presence of a linear time trend is proposed. The test is applicable in cross-correlated, heterogeneous panels and it can also be used when the integration order of innovations is unknown, by means of subsampling.

In the next chapter a cointegration test having asymptotic standard normal distribution and not requiring exogeneity assumptions is derived. In panels exhibiting cross-correlation or cointegration, individual test statistics are asymptotically independent, which leads to a panel test statistic robust to dependence across units.

The third chapter examines in an econometric context the simple idea of combining p-values from a series of statistical tests and improves its applicability in the presence of cross-correlation.

The last chapter applies recent panel techniques to OECD long-term interest rates and differentials thereof, finding only rather week evidence in favor of stationarity when allowing for cross-correlation.

Acknowledgements

The following is my heartfelt appreciation to those people, both past and present, who gave me the spirit, knowledge and insight to begin, conduct and complete this thesis.

First and foremost, I would like to acknowledge my adviser, Prof. Dr. Uwe Hassler. Throughout the course of my research he provided me with timely and valuable advice and excellent facilities. His courtesy, professionalism, and patience made working with him very rewarding and gratifying. I would like to thank my committee members for their constructive criticisms and useful suggestions and in particular, to my second examiner, Prof. Dr. Michael Binder. My colleagues, Dr. Matei Demetrescu and Vladimir Kuzin, provided me with useful discussions and a stimulating work environment.

I would also like to express gratitude towards the people who taught me mathematics: my father, Prof. Dr. Constantin Tarcolea, and my math teacher, Mr. Armand Martinov. My statistics knowledge is due to Prof. Dr. Virgil Craiu and Prof. Dr. Vasile Preda from University Bucharest. Two persons offered me guidance, support and advice in particular at the beginnings of my academic career: Prof. Dr. Hans-Dieter Heike and Prof. Dr. Mariana Craiu - I am deeply indebted to them.

My research took place within the project "Integration and Cointegration Analysis with Panel Data", financed by the Deutsche Forschungsgemeinschaft (German Research Foundation), to which I express my thanks.

Finally, I would like to thank my family, for their constant help and encouragement. I am also grateful to all my true and dear friends, because I know this research has profited from their support.

Table of Contents

List of Figures

List of Tables

Preface

Panel data has proven to be a valuable tool in describing, explaining, forecasting and also controlling the evolution of economic, but also social and technical processes. Panels obviously contain more information than single time series. Hence, they are used to gain power when testing economic hypotheses or to gain precision when estimating relevant parameters. However, macroeconomic panels pose particular difficulties, as will shortly be seen.

The first aspect to consider is that each of the single time series in a panel may be instationary. This affects the inference about the examined relationships between the observed macroeconomic variables. There are many forms of instationarity, and at least two should be mentioned: instationarity in the mean - in the most common form of linear trends - and instationarity in the variance, usually modelled by integrated processes. The latter are said to be driven by a stochastic trend. A process which is integrated of order 1 - I(1) - is stationary after building differences; equivalently, it is said to possess an autoregressive unit root.

Instationarity in form of trends (deterministic or stochastic) matters because there is a larger chance in this instationary context to observe spurious regressions than in the stationary framework. The question, whether there is a relationship among the observed series, turns into the question of whether there are common trends. For stochastic trends, this question can be answered with the help of cointegration analysis. The integrated time series are cointegrated if and only if there is a linear combination thereof that is stationary. Alternatively, cointegration can be explained through an error-correction mechanism, where the increments at time t are dependent on the deviations from the long term equilibrium at time $t-1$ in such a way that the balance is redressed. Cointegration describes hence a long term equilibrium. For deterministic trending, see the related notion of cotrending introduced by Hatanaka and Yamada (2003). Since many macroeconomic variables, such as interest rates, price levels or exchange rates, are considered as being integrated of order one, the empirical relevance of the problem is obvious.

On the other hand, the dependencies between the units of the panel can be a substantial disturbing factor in a panel analysis: for instance, it is difficult to decide if the statistical evidence in favour of a studied hypothesis arises because of the power gain of the panel test or if it is simply some result obtained by chance and influenced by the dependence of units. Such interdependencies are to be expected especially with regional data, where the observations of neighboring regions (countries) are usually high correlated.

Thus, tests in nonstationary panels are difficult to build because of cross-dependencies among units, which may lead to severe biases and serious complications with the non-standard asymptotics. This is best seen in the case of panel unit root testing. The first generation of unit root tests, see Levin, Lin and Chu (2002) or Im, Pesaran and Shin (2002), use an asymptotic theory typically built on the assumption of independent units, hypothesis not holding with macroeconomic panels. The second generation of tests tried to take this into account (for a survey, see Breitung and Pesaran, 2006). Two strategies were developed to this purpose. On the one hand, the conventional methods were adapted in such a way that they became robust to contemporary cross-correlations, see Chang (2002). An alternative approach resorted to generalized estimation methods (for instance Generalized Least-Squares), where the correlation matrix of the residuals is estimated in a first step and a transformation is carried out in the second step, transformation that can lead to an asymptotic efficient estimation, as in Breitung and Das (2005).

This thesis discussess and proposes new ways of dealing with precisely this category of possibly non-stationary and cross-dependent panels. The suggested procedures are carefully quantified by means of Monte Carlo experiments. Typical applications of the developed methods consist in multi-country studies, with several countries observed over a couple of decades. The empirical applications implemented here are the testing for trends in the investment share in European GDPs and the examination of OECD interest rates.

In the first chapter, a panel test for the presence of a linear time trend is proposed. It is important to have a test that can discern between a time trend and some random behavior in the presence of correlation. Testing for the presence of a linear trend is indeed a topic of great practical relevance, which has applications ranging from the technical analysis of financial markets to the prediction of global temperature. The proposed test is applicable in cross-correlated, heterogenous panels, as long as the number of units is smaller than the number of time observations. Further, it is shown how to apply the test when there is uncertainty about whether the innovations are stationary or integrated, by means of subsampling. Subsampling also allows for arbitrary number of units and works better in small samples than the asymptotic variant of the test. Finally, the method is applied to test whether the investment share in European GDPs has a trend or not.

In the next chapter, the issue of testing cointegration in cross-dependent panels is dealt with. A single equation, error correction model of the type employed by Banerjee, Dolado and Mestre (1998) is estimated using the non-linear instruments proposed by Chang (2002) for the augmented Dickey-Fuller test. When testing the null of no cointegration for a single unit, this yields a test statistic having asymptotic standard normal distribution without exogeneity assumptions, irrespective of the number of integrated covariates. In panels exhibiting cross-correlation or cointegration across units, individual test statistics are asymptotically independent, which leads to a panel test statistic robust to dependence across units.

The third chapter examines in an econometric context the simple idea of combining p-values from a series of statistical tests. The idea of combining significance of correlated statistics is taken up from biometry and medical science, so it also has a large field of applicability. Here are combined not the test statistics of single units (where the units can be mice as well as countries), but their corresponding p-values. The method is suitable for studies where the outcomes of the experiments are very different and hence impossible to combine directly. The original variant, the inverse normal method, requires independence of single test statistics in order to obtain asymptotic normality of the joint test statistic. This chapter discusses the modification due to Hartung (1999), which is designed to allow for a certain correlation matrix of the transformed p-values. Firstly, the modified inverse normal method is shown here to be valid with more general correlation matrices. Secondly, a necessary and sufficient condition for (asymptotic) normality is provided, using the copula approach. Thirdly, applications to panels of cross-correlated time series, stationary as well as integrated, are considered. This approach is not only adequate for many types of empirical problems, but also easy to implement and apply. A comparison with his direct competitors (which are implemented in EViews and do not allow for cross-correlation) shows that the proposed method performs well, even if some of the assumptions are violated.

The last chapter applies recent panel techniques to OECD interest rates and in particular to interest rate differentials. More precisely, panel unit root tests as well as procedures that combine individual significance levels to one joint p-value are employed. It turns out that empirical results largely depend on whether a homogeneous autoregressive structure is assumed for all countries or not. The assumption of independent units seems to be just as crucial. Only rather weak evidence in favour of stationary interest rate differentials is found when cross-correlation is allowed by using the modified inverse normal method.

Summing up, the new methods developed in this thesis should improve the reliability of panel studies with non-stationary, cross-correlated data and also provide answers to some empirical problems. As a matter of course, the results can be used to examine other economic hypotheses, provided that they imply long term relationships.

Chapter 1

Testing for Linear Trends in Dependent Heterogeneous Panels

1.1 Introduction

In many branches of applied science, testing for the presence of a linear trend is a topic of great practical relevance. Global warming can be mentioned as a prominent, although non-economic example. In applied economics, many studies have been centered on the verification (or rejection) of the Prebisch-Singer hypothesis: among others, Kim et al. (2003) and Bunzel and Vogelsang (2003). Of course, other issues of importance for policy makers can be quantified in terms of deterministic trends as well.

One aspect that has been extensively discussed in the context of testing for a time trend, see for instance Canjels and Watson (1997) or Sun and Pantula (1999), is whether the stochastic component is stationary or integrated. Vogelsang (1998) suggested a method that works in both situations; Bunzel and Vogelsang (2003) provide an alternative.

All these papers deal with single time series. Although Kim et al. (2003) consider more time series, they do not provide an overall test, but study each of them separately. Considering this, the panel test proposed here is a natural and necessary development.

In panels, especially when N is not small compared to T, things become more complicated than for single time series, due to cross-dependence. A similar problem appears in panel unit root testing, see Pesaran and Breitung (2006). Here, we model this dependence by means of cross-correlation. The proposed test statistic can be used either for short-memory ($I(0)$) or for integrated ($I(1)$) panels. We show how to use subsampling in order to encompass both possibilities for panel testing, as an alternative to the proposals for single time series by Vogelsang (1998) or Bunzel and Vogelsang (2003).

To the structure of the chapter: At first the model is described and a panel test statistic which tests the presence of a linear trend in either short-memory ($I(0)$) or in integrated ($I(1)$) panels is proposed. Then we show how to apply the test when a priori information about innovations is not available, by means of subsampling. In the end, the proposal is substantiated both by Monte Carlo simulations and an empirical application to the investment share of per capita GDP of 15 European countries.

1.2 Model and test

Let $y_{t,i}$, $t \in \{1, \ldots, T\}$, $i \in \{1, \ldots, N\}$ be our panel with N units, each with T time observations. Assume

$$y_{t,i} = \alpha_i + \beta_i t + x_{t,i},$$

where $\mathbf{x}_t = (x_{t,1}, \ldots, x_{t,N})'$ is a zero-mean $I(0)$ or $I(1)$ variable. Also, denote $\mathbf{y}_t = (y_{t,1}, \ldots, y_{t,N})'$.

In principle, testing for

$$\beta_i = 0 \text{ vs. } \beta_i \neq 0$$

in single units is not particularly difficult, even when $x_{t,i}$ is $I(1)$.[1] But in cross-dependent panels, this typically requires the non-parametric estimation of the long-run covariance matrix Ω of $\mathbf{x}_t$ (or of $\Delta\mathbf{x}_t$, respectively). Although the problem was studied before in detail by Andrews in his paper from 1991 (see also Newey and West, 1987), the larger N is, the more unreliable the estimation of Ω becomes.

The basic idea is to avoid estimation by using a possibly random matrix, but one "proportional" to Ω, so that Ω cancels out. This idea was suggested by Kiefer, Vogelsang and Bunzel (2000). For an application to testing for uncorrelatedness of dependent time series, see Lobato (2001). Using the same approach, Breitung (2002) shows how to modify variance-ratio-type tests to consistently test for unit roots and cointegration.

[1] As already mentioned, complications arise only when the order of integration is not known.

1.2.1 I(0) panels

Assume first:

$$T^{-0.5}\sum_{j=1}^{[sT]}\mathbf{x}_j \Rightarrow \Omega^{0.5}\mathbf{W}(s) \quad \text{for } s\in[0,1], \tag{1.1}$$

where "$\Rightarrow$" stands for weak convergence in a suitable metric space of cadlag random functions defined on $[0,1]$, $\mathbf{W}(s)$ is a vector of N independent standard Wiener processes and Ω is the long-run covariance matrix of the multivariate process $\mathbf{x}_t$. This requirement can be seen as the defining property of short-memory processes (cf. Lo, 1991).

We will use a demeaned statistic which reacts to the presence of a trend. Firstly, we will remove the mean from the data to obtain invariance with respect to α_i. With the demeaned data, we build partial sums,

$$\mathbf{S}_t=\sum_{j=1}^{t}\left(\mathbf{y}_j-\overline{\mathbf{y}}\right),$$

to obtain a statistic of the type used by Kwiatkowski et al. [KPSS] (1992):

$$\tau=\frac{1}{T^2}\sum_{t=1}^{T}\mathbf{S}_t'\mathbf{S}_t.$$

Under the null hypothesis, using the assumption on $\mathbf{x}_t$ and the Continuous Mapping Theorem, following distributional result arises imediately:

$$\tau \xrightarrow{d} \int_0^1 \mathbf{V}_\mu'(s)\,\Omega\mathbf{V}_\mu(s)\,ds,$$

where $\mathbf{V}_\mu$ is the so-called Brownian Bridge (see Kwiatkowski et al., 1992) and "$\xrightarrow{d}$" stands for convergence in distribution.

Remark 1.1 *If there is a deterministic trend in one of the units, it can be shown that*

$$\tau \xrightarrow{p} \infty \text{ with } \tau=O_p\left(T^3\right),$$

and the test rejects for large values.

Its asymptotic null distribution is not free of nuisance parameters, but a pivotal variant could be obtained by using transformed $\mathbf{S}_t$,

$$\mathbf{S}_t^{\perp}=\Omega^{-0.5}\mathbf{S}_t$$

to compute the corresponding statistic

$$\tau^{\perp} = \frac{1}{T^2} \sum_{t=1}^{T} \mathbf{S}_t^{\perp\prime} \mathbf{S}_t^{\perp},$$

for which the following relationship holds:

$$\tau^{\perp} \xrightarrow{d} \int_0^1 \mathbf{V}'_{\mu}(s) \, \mathbf{V}_{\mu}(s) \, ds.$$

However, we do not use this pivotal variant, which, as said, can be rather unreliable.

Instead, we use some function of the data "proportional" to Ω, but invariant to the presence of a linear trend, such that the nuisance matrix Ω cancels out under both null and alternative hypothesis. We denote by $\mathbf{y}_t^{\tau}$ the $\mathbf{y}_t$ after demeaning and detrending by OLS. The corresponding cumulative sums are defined accordingly:

$$\mathbf{S}_u^{\tau} = \sum_{j=1}^{u} \mathbf{y}_j^{\tau}.$$

We can now suggest the use of a detrended KPSS-type statistic:

$$\omega = \frac{1}{T^2} \sum_{u=1}^{T} \mathbf{S}_u^{\tau} \mathbf{S}_u^{\tau\prime} \tag{1.2}$$

and it results that

$$\omega \xrightarrow{d} \Omega^{0.5} \left(\int_0^1 \mathbf{V}_{\tau}(v) \, \mathbf{V}'_{\tau}(v) \, dv \right) \Omega^{0.5},$$

with $\mathbf{V}_{\tau}$ the second-level Brownian Bridge (see Kwiatkowski et al., 1992). The condition $N < T$ is required for invertibility of ω in (1.2).

Then, we define the test statistic for the presence of linear trends,

$$\tau^{*} = \sum_{t=1}^{T} \mathbf{S}'_t \left(\sum_{u=1}^{T} \mathbf{S}_u^{\tau} \mathbf{S}_u^{\tau\prime} \right)^{-1} \mathbf{S}_t. \tag{1.3}$$

Proposition 1.1 *Assuming (1.1) holds for $\mathbf{x}_t$, it follows for τ^* from (1.3) under the null hypothesis of no linear time trends*

$$\tau^{*} \xrightarrow{d} \int_0^1 \mathbf{V}'_{\mu}(s) \left(\int_0^1 \mathbf{V}_{\tau}(v) \, \mathbf{V}'_{\tau}(v) \, dv \right)^{-1} \mathbf{V}_{\mu}(s) \, ds$$

as $T \to \infty$.

Proof: Follows directly with the Continuous Mapping Theorem.

Remark 1.2 *Remark 1.1 can be shown to hold for τ^* as well, so the test is consistent against the alternative of linear time trend.*

The test rejects for too large values; the distribution, being non-standard, requires simulation of critical values. The simulations were carried out in GAUSS with 100000 replications and with $T = 1000$. The disturbances $x_{t,i}$ are generated as independent standard normal and we set $\alpha_i = \beta_i = 0$. We tabulate the critical values for $N \in \{1, 2, 5, 10, 15\}$ and the 1%, 5% and 10% significance levels, see Table 1.1.

Level	$N = 1$	$N = 2$	$N = 5$	$N = 10$	$N = 15$
1%	17.89	30.45	76.05	176.10	308.19
5%	9.39	18.13	51.01	128.09	233.93
10%	6.48	13.54	40.80	107.49	200.70

Table 1.1: Critical values for τ^* in the I(0) case

We now study the behaviour of the test in small samples. For our Monte Carlo study, we employ following data generating process:

$$\mathbf{x}_t = \mathbf{e}_t + \theta \mathbf{e}_{t-1}, \quad \mathbf{e}_t \sim iid\mathcal{N}\left(0, \Sigma\right),$$

with

$$\Sigma = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}.$$

Since many economic series are positively autocorrelated, we specify $\theta = 0.5$. Also, there are usually positive correlations between similar series across countries, so we set $\rho = 0.5$. The observed $\mathbf{y}_t$ is generated by adding a linear trend, with coefficient $\beta_i = \beta \in \{0, 0.01, 0.05, 0.1\}$. The rejection frequencies at the 5% level based on 5000 replications are given in Table 1.2, for sample sizes $T \in \{50, 100, 250\}$. The critical values from Table 1.1 were used.

The size is close enough to the nominal level only for very small N. For large N, the test is extremely liberal; this is especially striking for $T = 50$, but it improves with growing T. Convergence to the asymptotic behaviour

β	T	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	50	6.40	6.66	10.78	22.76	53.22
0	100	5.74	6.90	8.10	11.84	27.38
	250	4.92	5.40	5.94	7.22	15.68
	50	11.28	11.62	15.26	26.56	57.00
0.01	100	41.84	40.68	36.40	38.68	56.04
	250	100.00	100.00	99.96	100.00	100.00
	50	84.88	85.22	84.48	87.16	95.88
0.05	100	100.00	100.00	100.00	100.00	100.00
	250	100.00	100.00	100.00	100.00	100.00
	50	99.96	99.98	99.84	99.92	100.00
0.1	100	100.00	100.00	100.00	100.00	100.00
	250	100.00	100.00	100.00	100.00	100.00

Table 1.2: Size (5%) and power of test based on τ^* in the I(0) case

can be observed; the larger N is, the slower the convergence. Hence, it might be advisable to use exact finite-sample critical values when N is large. We however, favour a different solution, see subsection 1.2.3. In terms of power, the test behaves satisfactorily (however, given the size distortion, these rejection frequencies are not that meaningful).

1.2.2 I(1) panels

Should the stochastic components $x_{t,i}$ be integrated of order one (or, correspondingly, $\Delta \mathbf{x}_t$ have short memory), we assume:

$$T^{-0.5}\mathbf{x}_{[sT]} \Rightarrow \Omega^{0.5}\mathbf{W}(s), \quad \text{for } s \in [0,1]. \tag{1.4}$$

It turns out that the same test statistic τ^* has a well-defined limiting distribution under the null of no linear trends, see the proposition below. The test statistic thus remains

$$\tau^* = \sum_{t=1}^{T} \mathbf{S}_t' \left(\sum_{u=1}^{T} \mathbf{S}_u^\tau \mathbf{S}_u^{\tau\prime} \right)^{-1} \mathbf{S}_t.$$

Proposition 1.2 *Assuming (1.4) holds for $\mathbf{x}_t$, following asymptotic distribution results for τ^* under the null hypothesis of no linear time trends as $T \longrightarrow \infty$:*

$$\int_0^1 \left(\int_0^r \mathbf{W}_\mu(v)dv \right)' \left[\int_0^1 \left(\int_0^u \mathbf{W}_\tau(v)dv \right) \left(\int_0^u \mathbf{W}_\tau(v)dv \right)' du \right]^{-1} \left(\int_0^r \mathbf{W}_\mu(v)dv \right) dr,$$

where $\mathbf{W}_\mu$ is the demeaned Wiener process, $\mathbf{W}_\mu(v) = \mathbf{W}(v) - \int_0^1 \mathbf{W}(s)\,ds$, and $\mathbf{W}_\tau$ is the demeaned and detrended Wiener process (cf. Park and Phillips, 1988), $\mathbf{W}_\tau(v) = \mathbf{W}(v) + (6v-4)\int_0^1 \mathbf{W}(s)\,ds + (-12v+6)\int_0^1 s\mathbf{W}(s)\,ds$.

Proof: Follows directly with the Continuous Mapping Theorem.

Remark 1.3 *Under the alternative, the behaviour also changes compared to the $I(0)$ case: it is straightforward to show that*

$$\tau^* \xrightarrow{p} \infty, \text{ with } \tau^* = O_p(T)$$

as $T \to \infty$, so power reduction in comparison to the $I(0)$ case is to be expected.

As an alternative method, one might difference the series and test for a non-zero mean by a similar combination of KPSS-type statistics. However, this alternative test statistic would have a degenerate distribution if the series were $I(0)$, so we do not elaborate this topic here.

We have to give new critical values for the case of integrated disturbances, see Table 1.3.

Level	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
1%	351.53	1036.08	7568.4	48612.3	163848.8
5%	129.77	452.02	3934.5	29322.1	107222.5
10%	74.99	278.88	2771.4	22245.2	84373.4

Table 1.3: Critical values for τ^* in the I(1) case

Unfortunately, these differ strongly from those in the $I(0)$ case. Therefore, special care is to be taken if the integration order of the stochastic component $\mathbf{x}_t$ is not known, see the next subsection.

The corresponding Monte Carlo results (for which $\mathbf{x}_t$ is based on the same MA process as before, but using its partial sums) are provided in Table 1.4.

Compared to Table 1.2, the simulations provide good results regarding the size properties. The low power is determined by the fact that the small trend is "hidden" by the stochastic trend (note also the reduced convergence rate under the alternative hypothesis).

1.2.3 Unknown integration order

Since the test statistic τ^* has well-defined asymptotic null distributions under both $I(0)$ and $I(1)$ possibilities, while being consistent under the alternative

β	T	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	50	5.54	5.16	5.58	6.66	8.14
0	100	5.08	5.04	5.72	5.36	5.60
	250	5.56	4.28	5.10	5.30	4.84
	50	5.76	5.22	5.22	6.38	7.80
0.01	100	5.30	4.46	4.96	5.20	6.00
	250	5.44	5.18	5.60	5.16	5.34
	50	5.96	5.02	5.32	6.62	8.78
0.05	100	5.22	5.70	5.68	5.14	5.42
	250	6.64	6.36	6.40	6.12	5.20
	50	6.42	6.10	6.40	7.28	8.72
0.1	100	8.50	6.88	6.38	6.52	5.96
	250	11.58	10.86	9.46	6.90	7.38

Table 1.4: Size (5%) and power of test based on τ^* in the I(1) case

of a linear trend, the subsampling method, proposed by Politis and Romano (1994), works in both $I(0)$ and $I(1)$ cases.

The idea of the subsampling method is to approximate the sampling distribution of a test statistic by recomputing it on subsamples of smaller size of the observed data. One uses blocks of size l of consecutive observations in order to preserve within each block the dependence structure of the underlying model. Note that there are $M = T - l + 1$ such blocks. Only a very weak assumption on l will be required; typically, $l/T \to 0$ and $l \to \infty$ as $T \to \infty$.

Compared to the bootstrap methods, for which much work has to be done in order to demonstrate their validity in a given situation, subsampling only requires the existence of a nondegenerate limiting distribution of the respective statistic and some bound on the serial dependence of the underyling sequence. In the context of time series this can be for instance a mixing condition, see Politis, Romano and Wolf (1999, p. 70).

In the panel situation, the series modelling the disturbances must satisfy a mixing condition corresponding to the multivariate framework. Together with a moment condition, this implies weak convergence to Brownian Motion (see, for instance, Herrndorf, 1984), which is required for the test to have a valid asymptotic distribution. Of course, in the case of integrated disturbances, the mixing condition does not hold. One may still use subsampling, since, when *the differences* are mixing, subsampling these would be justified. Because the test statistic is invariant w.r.t. non-zero mean, subsampling differences and building their partial sums is equivalent to subsampling the disturbances themselves. Hence, one may subsample the data in both $I(0)$ and $I(1)$ cases.

Similarly, to simplify the computation, we suggest as panel test statistic the sum of single test statistics:

$$\tau^S = \sum_{i=1}^{N} \tau_i^*,$$

where τ_i^* are the univariate analogues to (1.3). The distribution of this modified τ^S is then subsampled.

Summing single test statistics removes the requirement that N be smaller than T. The distribution of τ^S depends indeed on Ω, but, by using subsampled critical values (as described below), the test becomes invariant to Ω without requiring any kind of orthogonalization. More importantly, cointegration across units is now allowed for $I(1)$ series by the same argument.

We detail here the used subsampling procedure: First build $M = T - l + 1$ overlapping blocks of length l, the first one being $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_l\}$ and the last one being $\{\mathbf{y}_{T-l+1}, \mathbf{y}_{T-l+2}, ..., \mathbf{y}_T\}$. Let B_m be a generic block (of size l) of the consecutive data $\{\mathbf{y}_m, ..., \mathbf{y}_{m+l-1}\}$, $m = 1, 2, ..., M$; we compute for each of the blocks the corresponding τ_m^S statistic and then use the M ordered realizations τ_m^S in order to estimate the distribution function. The critical value at level α is then $\tau^S_{[(1-\alpha)M]}$.

The optimal choice of the block size l can be data-driven if the integration order of the series is known. Since our assumption is that we do not know the true integration order, we have to run a Monte Carlo experiment to determine optimal values for l. We must note here the fact that, although l converges to infinity together with T, the rate of convergence of τ^S is slower than in the parametric case and hence the power of the subsampled test is expected to be lower.

The results for the Monte Carlo simulations for $T = 50, 100$ and 250 are given in Tables 1.5, 1.6 and 1.7, respectively. A few words to the Monte Carlo setup: since our goal is only to determine an optimal l, and not to test, we simplify the framework as much as possible. So we set $\theta = \rho = 0$, which also ensures "neutrality" of the Monte Carlo experiment, and let β take only two possible values, $\beta \in \{0, 0.1\}$. The simulations are carried out with 5000 replications and each table contains the rejection frequencies at the 5% level for both $I(0)$ and $I(1)$ cases.

We choose the values of l which realize a reasonable trade-off between correct size and power. In Tables 1.5, 1.6 and 1.7, the bolded cases represent our choice of optimal l. From the simulations results we see that the optimal

l	β	$I(\cdot)$	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	0	0	3.52	1.82	0.86	0.26	0.18
6	0	1	10.02	10.80	10.90	11.10	10.50
	0.1	0	99.88	99.96	100.00	100.00	100.00
	0.1	1	11.92	15.74	17.08	19.36	22.82
	0	0	**5.94**	5.20	3.08	1.68	1.70
8	0	1	**10.92**	10.42	12.00	13.88	14.52
	0.1	0	**99.88**	99.98	100.00	100.00	100.00
	0.1	1	**14.54**	15.10	19.78	24.24	27.58
	0	0	7.52	**6.50**	5.20	4.32	3.32
10	0	1	10.58	**10.80**	12.44	13.06	15.06
	0.1	0	99.88	**100.00**	100.00	100.00	100.00
	0.1	1	14.50	**16.50**	18.58	23.66	27.06
	0	0	6.90	6.00	**4.74**	3.64	3.40
12	0	1	8.92	9.54	**10.36**	10.68	10.88
	0.1	0	99.40	99.98	**100.00**	100.00	100.00
	0.1	1	10.98	13.06	**14.88**	18.56	21.02
	0	0	7.52	6.24	6.02	**5.40**	**4.52**
14	0	1	9.26	9.90	9.38	**10.46**	**11.96**
	0.1	0	99.00	99.94	100.00	**100.00**	**100.00**
	0.1	1	11.10	12.56	13.70	**18.12**	**19.36**
	0	0	8.52	7.14	6.48	6.56	6.42
16	0	1	9.60	9.98	10.86	11.46	11.52
	0.1	0	98.44	99.90	100.00	100.00	100.00
	0.1	1	12.20	12.52	14.12	17.78	20.34

Table 1.5: Size (5%) and power of test based on τ^S as function of block length, T=50

l	β	$I(\cdot)$	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	0	0	**5.62**	4.90	4.10	2.94	2.94
12	0	1	**8.08**	8.64	9.42	9.98	10.86
	0.1	0	**100.00**	100.00	100.00	100.00	100.00
	0.1	1	**13.18**	15.94	21.72	29.42	34.88
	0	0	6.56	**5.60**	5.46	4.26	4.38
14	0	1	8.66	**8.56**	10.44	10.50	10.22
	0.1	0	100.00	**100.00**	100.00	100.00	100.00
	0.1	1	13.48	**16.68**	21.14	26.64	33.32
	0	0	6.78	6.88	**5.46**	5.46	5.42
16	0	1	9.46	9.10	**9.14**	9.52	9.80
	0.1	0	100.00	100.00	**100.00**	100.00	100.00
	0.1	1	14.34	16.70	**20.50**	26.00	32.26
	0	0	7.66	7.24	6.18	**5.62**	**4.98**
18	0	1	9.12	9.40	9.04	**9.82**	**9.62**
	0.1	0	100.00	100.00	100.00	**100.00**	**100.00**
	0.1	1	14.60	15.44	19.94	**26.52**	**30.84**
	0	0	7.98	7.88	7.46	6.02	6.18
20	0	1	9.92	9.02	8.80	9.74	10.38
	0.1	0	100.00	100.00	100.00	100.00	100.00
	0.1	1	13.80	16.42	20.88	24.70	29.32

Table 1.6: Size (5%) and power of test based on τ^S as function of block length, T=100

block length also depends on N and that it would be appropriate to choose larger l for larger N. This happens because the overall variability in data increases with growing N, except for the case when T is large and the number of observations compensates the increase in data variability.

Finally, we provide the Monte Carlo simulation results for the behaviour of test itself in Tables 1.8 and 1.9, which are the "subsampling equivalents" of Tables 1.2 and 1.4.

Compared to Table 1.2, subsampling in the $I(0)$ case stabilizes the size (the test is even somewhat conservative). This comes at a cost of power losses, which is a common problem in the subsampling framework. However, one should note that, for the subsampled version, power increases with N, which does not seem to be the case with the parametric version of the test.

In the $I(1)$ case, subsampling leads to an oversized test for small number of units. For a large number of units, the size remains close to the nominal level, which was not the case with the parametric version of the test. Furthermore, the power gains over Table 1.4 are significant for $N \geq 2$. Again, power increases with growing N.

l	β	$I(\cdot)$	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	0	0	5.40	4.98	4.88	4.18	3.74
18	0	1	6.78	6.64	7.20	8.14	8.04
	0.1	0	100.00	100.00	100.00	100.00	100.00
	0.1	1	19.00	24.02	37.44	52.24	64.62
	0	0	**5.36**	**5.12**	**5.04**	4.56	4.30
20	0	1	**6.58**	**6.58**	**7.36**	7.86	7.78
	0.1	0	**100.00**	**100.00**	**100.00**	100.00	100.00
	0.1	1	**20.32**	**23.46**	**35.76**	49.88	63.42
	0	0	5.74	5.92	5.56	**4.48**	**4.58**
22	0	1	7.14	6.70	7.54	**7.44**	**7.52**
	0.1	0	100.00	100.00	100.00	**100.00**	**100.00**
	0.1	1	19.20	23.60	35.30	**50.38**	**60.86**
	0	0	6.16	5.94	5.50	5.32	4.94
24	0	1	7.26	7.46	7.78	7.32	7.34
	0.1	0	100.00	100.00	100.00	100.00	100.00
	0.1	1	16.38	19.46	28.22	39.16	47.18

Table 1.7: Size (5%) and power of test based on τ^S as function of block length, T=250

To summarize these results, our advice to practitioners is to use the subsampled test, since it works better in most cases and does not impose any conditions on N and T; additionally, it provides trustworthy results in cases when the assumption about the integration order is wrong or not certain.

1.3 Investment shares of per capita GDP

As an empirical application, we examine the investment shares of per capita GDP, since investment shares are a relevant measure for an economy. Even the simplest endogenous growth model incorporates this quantity as a key variable. Along this line, Levine and Renelt (1992) find investment shares to correlate with growth and with the ratio of international trade to GDP. A more recent contribution in this field of research is due to Madsen (2002), who analyzes the interdependence between economic growth and investment.

We employ annual data from the Penn World Table (Heston, Summers and Aten, 2002). There is strong reason to assume the investment shares of the GDP not to be integrated and we use therefore critical values from Table 1.1. More precisely, we examine 15 European economies, members of the European Union (the former EU 15)[2]: Austria, Belgium, Denmark, Finland, France,

[2]Eight of the ten omitted new EU members, being during the 70's and the 80's members of the communist block, exhibit data that is not comparable to the EU 15.

β	T	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	50	4.20	3.60	4.00	4.60	3.90
0	100	2.60	3.00	3.40	3.60	4.60
	250	3.30	3.00	2.80	4.30	2.80
	50	6.40	6.80	6.80	8.50	8.20
0.01	100	27.60	33.60	49.00	52.40	57.00
	250	99.80	100.00	100.00	100.00	100.00
	50	56.20	71.70	79.70	87.20	88.10
0.05	100	99.90	100.00	100.00	100.00	100.00
	250	100.00	100.00	100.00	100.00	100.00
	50	94.80	98.20	99.60	99.40	99.70
0.1	100	100.00	100.00	100.00	100.00	100.00
	250	100.00	100.00	100.00	100.00	100.00

Table 1.8: Size (5%) and power of test based on τ^S in the I(0) case, l as bolded in Tables 1.5, 1.6 and 1.7.

β	T	$N=1$	$N=2$	$N=5$	$N=10$	$N=15$
	50	6.00	7.90	7.60	7.20	6.50
0	100	6.90	7.00	7.90	8.70	6.80
	250	5.00	6.40	7.60	6.20	5.60
	50	9.40	9.10	5.60	7.20	5.60
0.01	100	6.00	6.40	7.00	8.60	6.50
	250	6.80	7.20	6.60	6.60	5.00
	50	6.80	8.60	7.00	7.50	7.00
0.05	100	6.30	6.80	8.00	10.00	9.00
	250	6.50	8.20	9.40	9.20	9.60
	50	7.40	10.60	8.80	8.00	10.30
0.1	100	9.40	9.10	10.70	11.20	11.40
	250	9.80	14.00	14.70	15.50	19.20

Table 1.9: Size (5%) and power of test based on τ^S in the I(1) case, l as bolded in Tables 1.5, 1.6 and 1.7.

Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and the United Kingdom. The used data, covering the period from 1970 to 2000, is plotted in Figure 1.1.

The investment shares have a visible downward time trend in all countries except for Ireland, Luxembourg, Portugal and the UK, but the degree of co-movement seems considerable. The panel test statistic has a value of 4338.27, which is highly significant even at the 1% level. Recall, however, that the test based on asymptotic critical values is extremely oversized for $N = 15$. If the data were integrated, this would not be significant anymore, not even at the 10% level. In this case, however, the co-movement would most likely appear due to cointegration, case in which the critical values in Table 1.3 would not be valid anymore.

To make sure the finding of a trend is substantiated, we apply the subsampling procedure, first at the country level and then for the whole panel. The simulations to obtain an optimal l are repeated for $T = 30$ and $N = 1$, as well as $N = 15$. For the single tests, we find significant trends only in Austria, Finland, France, Italy and Sweden, which is somewhat surprising, given the graphical evidence. In this case, $l = 6$ is the optimal block length for which the test keeps its size, while the power properties are satisfactory.

Returning to panels, in the $I(0)$ case the size properties are good for $l = 9$ and $l = 10$; the test is undersized for $l = 11$, while loosing power at the same time. Similarly to Table 1.5, the test is oversized (around 10% for the 5% significance level) in the $I(1)$ case; $l = 10$ or $l = 11$ seem here to be the optimal choices. The power properties are improved compared to the single tests, as expected.[3]

To the test results: in this application, the importance of l becomes obvious at the panel level. If choosing $l = 9$, the subsampled test rejects at the 5% level. If choosing $l = 10$ or $l = 11$, the subsampled test only rejects at the 10% level.

To sum up, there is statistical evidence that investment activity has been reducing over the past 30 years in the EU area. One explanation that seems plausible is that the accelerated development of new technologies in the past 40 years reduced the level of investment necessary to keep growth at a given level.

[3]The complete results are available upon request from the author.

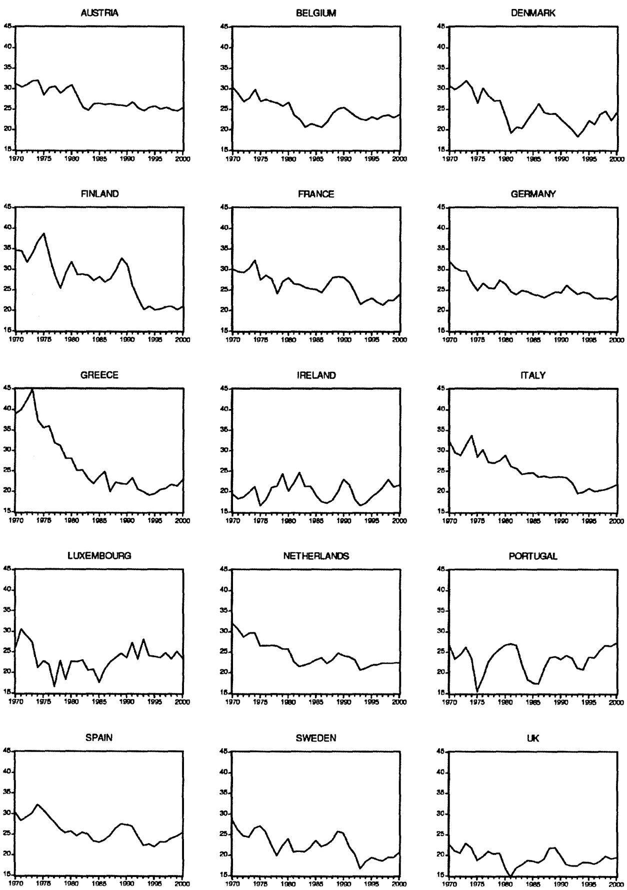

Figure 1.1: Investment shares of per capita GDP in EU15

1.4 Conclusions

We study a panel test for the presence of linear trends. The test works for cross-correlated, heterogeneous panels which are either stationary or integrated, as long as $N < T$.

Further, it is shown how to use subsampling to apply the test when the order of integration is not known. Subsampling also allows for arbitrary N and improves the small sample behaviour over the asymptotic approximation in terms of both size and power. Hence, the subsampled version of our test should be preferred in practical applications.

In an empirical application, it is found that the investment share of per capita GDP has been declining since 1970 in the EU area as a whole; for individual countries, the evidence is mixed. This result also emphasizes the importance of panel tests like the one proposed here.

Chapter 2

Panel Cointegration Testing using Nonlinear Instruments

2.1 Motivation

Using panel data is generally viewed as a method of gaining power when testing economic hypotheses and estimating relevant parameters. Although panels obviously contain more information than single time series, tests in nonstationary panels are difficult to build because of cross-correlation among units, which may lead to serious biases, see e.g. O'Connell (1998). Moreover, cointegration across units was also found to distort panel test statistics (Banerjee, Marcellino and Osbat, 2004). Then, one must either account for such cross-unit dependence, or use methods that are robust to it.

Popular tests for panel cointegration, like those of Kao (1999), McCoskey and Kao (1998) or Pedroni (2004), do not cope well with cross-unit dependence. One obvious fix is to bootstrap the respective test statistics, e.g. in the manner of Maddala and Wu (1999). This, however, comes at a cost: bootstrapping time series is notoriously difficult. Factor models, as advocated by Bai and Ng (2004), are able to deal with cross-unit dependence as well.

Here, we take a different approach, by adopting the single equation framework advocated by Banerjee, Dolado and Mestre (1998). Instead of using OLS estimation and testing, as Banerjee, Dolado and Mestre (1998) did, we suggest to employ nonlinear instrumental variables [NIV] of the kind Chang (2002) used in estimating the augmented Dickey-Fuller test regression.

Our main contribution is to analyze the asymptotic behavior of the testing procedure proposed above. For a single unit, we show this NIV cointegration test to have asymptotic standard normal distribution, even when regressors are not weakly exogenous. In a panel context, individual statistics are shown to be asymptotically independent in the presence of cross-unit correlation or cross-unit cointegration. This leads to a panel test statistic not affected by cross-unit dependence; moreover, standard limiting distributions result without N asymptotics. Work in progress by Chang (2005) looks into a NIV estimation procedure for residual-based cointegration tests.

The chapter is structured as follows. First, we describe the model we work with. Section 2.3 analyzes the NIV cointegration test in the error correction framework and its extension to panels exhibiting cross-dependence. Monte Carlo evidence on small-sample behavior is given in Section 2.4. The main findings and comments are summarized in the final section.

2.2 Model and assumptions

To reduce notational burden, we first study a single unit. The assumptions for the panel case are discussed in Section 2.3.2.

We assume the data to be modelled by an integrated vector autoregressive process of order $p+1$ with $K+1$ components, $K \geq 1$. This process is possibly cointegrated:

Assumption 2.1 *Let the observed data for one unit be generated as follows:*

$$\Delta \mathbf{w}_t = \Pi \mathbf{w}_{t-1} + \sum_{i=1}^{p} A_i \Delta \mathbf{w}_{t-i} + \boldsymbol{\epsilon}_t,$$

where $\mathbf{w}_t' = (w_{1,t}, \ldots, w_{K+1,t})$ *and* $\mathbf{w}_0 = \mathbf{0}$.

Denote with r the rank of Π. Under no cointegration, one has $\Pi = \mathbf{0}$, or $r = 0$. Under cointegration, one has $0 < r < K+1$ and the known factorization of Π, $\Pi = \alpha_r \beta_r'$, in two $(K+1) \times r$ matrices of adjustment speed coefficients and of parameters of the long-run relations, respectively, holds for $r > 0$. Under the alternative hypothesis of cointegration, we assume $r = 1$. This is needed to motivate the test statistic, see Banerjee, Dolado and Mestre (1998). The following assumption guarantees that the process $\mathbf{w}_t$ either follows a stable vector autoregressive process in differences (no cointegration), or, when $\Pi \neq 0$, I(2) processes are avoided.

Assumption 2.2 *Let the roots of the characteristic polynomial associated to* $\mathbf{w}_t$ *defined in Assumption 2.1 be either 1 or have absolute values larger than*

1. Further, if $\Pi = \mathbf{0}$*, let* $\det\left(\alpha'_{r\perp}\left(I - \sum_{i=1}^{p} A_i\right)\beta_{r\perp}\right) \neq 0$*, where* $\alpha_{r\perp}$ *and* $\beta_{r\perp}$ *are the orthogonal complements of* α_r *and* β_r *w.r.t.* $\mathbb{R}^{K+1}$*.*

Note that no additional restriction is imposed upon the elements of the matrices A_i or on Π. Also, the innovations $\boldsymbol{\epsilon}_t$ are allowed to correlate:

Assumption 2.3 *Let* $\boldsymbol{\epsilon}_t$ *obey following conditions:* $\boldsymbol{\epsilon}_t \sim iid\,(0, \Sigma)$*, with* Σ *any symmetric, positive definite* $(K+1) \times (K+1)$ *matrix. Furthermore, assume* $\exists l > 2$ *such that* $E\left\|\boldsymbol{\epsilon}_t\right\|^l < \infty$ *and let* $\boldsymbol{\epsilon}_t$ *have absolutely continuous distribution function (w.r.t. Lebesgue measure); let also* $\exists s > 0$ *such that* $\phi(\boldsymbol{\lambda}) = o\left(\|\boldsymbol{\lambda}\|^s\right)$*, where* ϕ *is the corresponding characteristic function.*

Assumptions 2.2 and 2.3 together allow for lack of exogeneity. This is an important aspect, since the weak exogeneity assumption is the main source of criticism to the single equation approach. The conditions in Assumption 2.3 are stronger than the typical sets of assumptions under which an invariance principle for cumulated innovations holds, but are needed in order to establish asymptotic behavior of regularly integrable transformations of integrated processes, see Park and Phillips (1999, 2001), as well as Chang, Park and Phillips (2001) and de Jong and Wang (2005). It is this framework that leads to robustness against cross-unit dependence in panels (see Subsection 2.3.2).

Denote $\mathbf{w}_{e,t} = (w_{2,t}, \ldots, w_{K+1,t})'$. In the single equation framework, the error correction representation can be then written as follows:

$$\begin{aligned} \Delta w_{1,t} &= \alpha\left(w_{1,t-1} + \boldsymbol{\theta}'\mathbf{w}_{e,t-1}\right) + \delta\left(L\right)\Delta w_{1,t-1} + \boldsymbol{\gamma}'(L)\Delta\mathbf{w}_{e,t-1} + \varepsilon_t, \quad (2.1) \\ \Delta\mathbf{w}_{e,t} &= \boldsymbol{\alpha}_e\left(w_{1,t-1} + \boldsymbol{\theta}'\mathbf{w}_{e,t-1}\right) + \boldsymbol{\delta}_e(L)\Delta w_{1,t-1} + \Gamma(L)\Delta\mathbf{w}_{e,t-1} + \boldsymbol{\nu}_t, \quad (2.2) \end{aligned}$$

for $t = 1, 2, \ldots, T$, where $\alpha \in \mathbb{R}$, $\boldsymbol{\alpha}_e \in \mathbb{R}^K$, $\boldsymbol{\theta} \in \mathbb{R}^K$, the respective lag polynomials and the innovations ε_t and $\boldsymbol{\nu}_t$ are defined implicitly from Assumption 2.1. Not including contemporaneous differences $\Delta\mathbf{w}_{e,t}$ in Equation (2.1), as Banerjee, Dolado and Mestre (1998) do, is compensated by having allowed for correlated innovations. Equation (2.1) can be transformed to contain only levels of the integrated variables $w_{1,t}$ and $\mathbf{w}_{e,t}$, if one is interested in the pure ADL representation of the model.

The null hypothesis in the single equation framework is $\alpha = 0$; in case of weak exogeneity, $\boldsymbol{\alpha}_e = \mathbf{0}$ is an implicit additional assumption. However, we wish to extend the types of endogeneity our test copes with, and thus explicitly allow the error correction to affect Equations (2.2). Hence, we shall examine following null:

Null hypothesis: $\alpha = 0$.

Note that, when allowing for error correction to affect the other components of $\mathbf{w}_t$, the null of the test is actually absence of error correction in the studied equation and not lack of cointegration between $w_{1,t}$ and $\mathbf{w}_{e,t}$. This attribute is common to all approaches based on a single equation and not specific to our test. See Remark 2.6 further below for a simple solution to this problem.

Under the alternative, α needs to be negative if error-correction is present only in Equation (2.1). Otherwise, α may also be positive (see Johansen, 1995, p. 54, for an example). Thus, our alternative hypothesis is as follows:

Alternative hypothesis: $\alpha \neq 0$.

The basic idea of our test is to replace OLS estimation of the test Equation (2.1) with instrumental estimation using regularly integrable transformations. Specifically, $F(w_{1,t-1})$ is used as instrument for $w_{1,t-1}$, where $F(\cdot)$ is restricted as follows:

Assumption 2.4 *Let $F(\cdot)$ be continuous on $\mathbb{R}$ with $\int_{-\infty}^{\infty} xF(x)dx$ finite and non-zero. Assume further that $|F(\cdot)|$ is bounded by a function $R(\cdot)$, where $R(\cdot)$ is integrable, continuous on $\mathbb{R}$ and monotone on $(-\infty, 0)$ and $(0, \infty)$.*

In what concerns the other integrated regressors, two possibilities arise. First, we may take them as instruments for themselves. Second, we may take regularly integrable transformations as instruments. We shall call the first case "partial instrumentalization", and the second will be denoted as "complete instrumentalization". However, Monte Carlo experiments (see Section 2.4) show the completely instrumentalized test to have very low power. Therefore, this paper focusses on partial instrumentalization.

The described data generating process does not exhibit deterministic components. These will be dealt with in Section 2.3.1.

2.3 Asymptotic results

As already mentioned, we first deal with the case of a single unit. The panel case is discussed in the second subsection.

2.3.1 Single unit test

The test regression (2.1) is reparameterized to match the usual notation of the single equation framework. Defining

$$\begin{aligned} \mathbf{x}_{t-1}' &= \left(\Delta w_{1,t-1}, \ldots, \Delta w_{1,t-p}, \Delta \mathbf{w}_{e,t-1}', \ldots, \Delta \mathbf{w}_{e,t-p}', \mathbf{w}_{e,t-1}'\right) \\ &= \left(\mathbf{x}_{t-1,0}', \mathbf{x}_{t-1,1}'\right), \end{aligned}$$

the I(1) variables ($\mathbf{x}_{t-1,1} = \mathbf{w}_{e,t-1}$) are separated from the I(0) ones ($\mathbf{x}_{t-1,0}$). The single-equation model becomes with $y_t = w_{1,t}$

$$\Delta y_t = \alpha y_{t-1} + \boldsymbol{\beta}' \mathbf{x}_{t-1} + \varepsilon_t, \tag{2.3}$$

with a new parameter vector $\boldsymbol{\beta}$,

$$\boldsymbol{\beta}' = \left(\delta_1, \ldots, \delta_p, \boldsymbol{\gamma}_1', \ldots, \boldsymbol{\gamma}_p', \alpha \boldsymbol{\theta}'\right),$$

where δ_i and $\boldsymbol{\gamma}_i$, $i = 1, 2, \ldots, p$, are the respective coefficients of the lag polynomials from Equation (2.1). It is convenient to write $\boldsymbol{\beta}' = (\boldsymbol{\beta}_0', \boldsymbol{\beta}_1')$, with $\boldsymbol{\beta}_0' = \left(\delta_1, \ldots, \delta_p, \boldsymbol{\gamma}_1', \ldots, \boldsymbol{\gamma}_p'\right)$ and $\boldsymbol{\beta}_1' = \alpha \boldsymbol{\theta}'$, in accordance to $\mathbf{x}_{t-1}' = \left(\mathbf{x}_{t-1,0}', \mathbf{x}_{t-1,1}'\right)$.

The t statistic of the estimated parameter $\widehat{\alpha}$ remains the natural choice as a test statistic for the null $\alpha = 0$, even with instrumental estimation. Note that, under the null hypothesis $\alpha = 0$, it holds $\boldsymbol{\beta}_1 = \mathbf{0}$. The parameter vector $\boldsymbol{\theta}$ is not identified under the null hypothesis, but α and $\boldsymbol{\beta}_1$ are identified under both null and alternative hypothesis. Hence, there is no impediment in testing this way.

For the case of partial instrumentalization, one obtains with the help of standard regression algebra,

$$\widehat{\alpha} - \alpha = Q^{-1} M,$$

where

$$M = \sum_{t=1}^{T} F\left(y_{t-1}\right) \varepsilon_t - \sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}_{t-1}' \left(\sum_{t=1}^{T} \mathbf{x}_{t-1} \mathbf{x}_{t-1}'\right)^{-1} \sum_{t=1}^{T} \mathbf{x}_{t-1} \varepsilon_t$$

and

$$Q = \sum_{t=1}^{T} F\left(y_{t-1}\right) y_{t-1} - \sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}_{t-1}' \left(\sum_{t=1}^{T} \mathbf{x}_{t-1} \mathbf{x}_{t-1}'\right)^{-1} \sum_{t=1}^{T} \mathbf{x}_{t-1} y_{t-1}.$$

For the t statistic, it holds under the null hypothesis $\alpha = 0$

$$t_{\widehat{\alpha}} = \frac{\widehat{\alpha}}{\widehat{\sigma}_{\widehat{\alpha}}},$$

with $\widehat{\sigma}_{\alpha}$ the estimated standard deviation of $\widehat{\alpha}$:

$$\widehat{\sigma}^2_{\widehat{\alpha}} = \widehat{\sigma}^2_{\varepsilon} Q^{-2} P,$$

where $\widehat{\sigma}^2_{\varepsilon}$ is a consistent estimator of the residual variance, $\sigma^2_{\varepsilon} = Var(\varepsilon_t)$, and

$$P = \sum_{t=1}^{T} F\left(y_{t-1}\right)^2 - \sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}'_{t-1} \left(\sum_{t=1}^{T} \mathbf{x}_{t-1} \mathbf{x}'_{t-1}\right)^{-1} \sum_{t=1}^{T} \mathbf{x}_{t-1} F\left(y_{t-1}\right),$$

which leads to

$$t_{\widehat{\alpha}} = \frac{M}{\widehat{\sigma}_{\varepsilon}\sqrt{P}}.$$

In the following, it will be more convenient to study the pivotal statistic

$$t^* = \frac{\widehat{\alpha} - \alpha}{\widehat{\sigma}_{\widehat{\alpha}}} = \frac{M}{\widehat{\sigma}_{\varepsilon}\sqrt{P}},$$

irrespective of the value of α.

The following proposition establishes convergence properties of the NIV estimators (see the Appendix for proof details).

Proposition 2.1 *For partially instrumentalized NIV estimation of test equation (2.3), it holds as $T \to \infty$ under Assumptions 2.1 through 2.4 and $\alpha = 0$ that*

a) if $\boldsymbol{\alpha}_e = \mathbf{0}$

$$\widehat{\alpha} - \alpha = O_p\left(T^{-0.75}\right), \tag{2.4}$$

$$\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0 = O_p\left(T^{-0.25}\right) \tag{2.5}$$

and

$$\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 = O_p\left(T^{-0.75}\right) \tag{2.6}$$

b) if $\boldsymbol{\alpha}_e \neq \mathbf{0}$,

$$\widehat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta}_0 = O_p\left(T^{-0.25}\right).$$

However, the convergence rates of $\widehat{\alpha}$ or $\widehat{\boldsymbol{\beta}}_1$ to the true values α and $\boldsymbol{\beta}_1$ can be $T^{0.25}$, depending on the cointegrating vector $(1, \boldsymbol{\theta}')'$.

Proof: *See the Appendix.*

This behavior of the estimators associated to the integrated regressors appears under both the null and the alternative hypothesis, since we allowed for lack of exogeneity in form of error-correction in the other equations of the error correction model besides the test Equation (2.1). The direct effect on the test is that one should not use NIV residuals when computing the residual variance, since convergence rates of the estimators may not be high enough to ensure consistent residuals, given that the regressors which $\widehat{\alpha}$ and $\widehat{\boldsymbol{\beta}}_1$ are associated to are integrated. This seems to be the case even when y_{t-1} and $\mathbf{x}'_{t-1,1}$ do not cointegrate (and the convergence rates of the estimators are high enough); a Monte Carlo experiment for the simplest case of two independent random walks with no deterministic components or short-run dynamics suggests that the behavior of the estimators is not very reliable and the variance of the test statistic under the null hypothesis is lower than unity by a factor of up to 1.4 for sample sizes between $T = 100$ and $T = 500$. One should resort to alternative residual variance estimators, such as using residuals from the OLS estimation of the model, which is employed throughout this paper.

Fortunately, the behavior of the pivotal statistic t^* is only indirectly affected by this, by means of the residual variance estimator. The following lemma eases the discussion of the suggested test. Its proof can be found in the Appendix as well.

Lemma 2.1 *If using a consistent residual variance estimator, it holds under the assumptions of Proposition 2.1 that*

$$t^* = \frac{T^{-0.25} \sum_{t=1}^{T} F(y_{t-1}) \varepsilon_t}{\sigma_\varepsilon \sqrt{T^{-0.5} \sum_{t=1}^{T} F(y_{t-1})^2}} + o_p(1) \ ,$$

regardless of whether $\boldsymbol{\alpha}_e = \mathbf{0}$ *or* $\boldsymbol{\alpha}_e \neq \mathbf{0}$.

Proof: *See the Appendix.*

Remark 2.1 *It is now clear why no restrictive weak exogeneity assumptions have to be made. This is because no regressor except the lagged dependent variable* y_{t-1} *influences the test statistic for large* T *- as can be seen from Lemma 2.1. In contrast to that, OLS estimation of the test equation requires either weak exogeneity or inclusion of leads to account for second-order bias (see Banerjee, Dolado and Mestre, 1998).*

The following proposition summarizes the asymptotic behavior of the proposed test statistic under both null and alternative hypothesis.

Proposition 2.2 *Under the assumptions of Lemma 2.1, it holds as $T \to \infty$:*

a) if $\alpha = 0$, then

$$t_{\widehat{\alpha}} \xrightarrow{d} \mathcal{N}(0,1);$$

b) if $\alpha \neq 0$, then

$$|t_{\widehat{\alpha}}| \xrightarrow{p} \infty.$$

Proof: *See the Appendix.*

Remark 2.2 *Confidence intervals for the parameter α are straightforward to build, since the pivotal statistic t^* has asymptotic standard normal distribution whatever the true value of α is, as long as a consistent residual variance estimator is used.*

Remark 2.3 *Following Demetrescu (2006), Proposition 2.2 can also be established for the case where $\Delta \mathbf{w}_t$ follows a general linear process with a weak summability condition, when an autoregressive approximation of order growing to infinity, but slower than T, is used. This rate should be $o(T^{0.25})$, but not slowly varying at infinity, see Demetrescu (2006) for further details.*

Remark 2.4 *In practice, the order of the autoregressive process capturing short-run dynamics is of course not known. Due to Lemma 2.1, data-driven lag order choice (such as sequential significance testing of the autoregressive parameters) will have no asymptotic effect on the test statistic; see Demetrescu, Kuzin and Hassler (2006) for a situation where this is not the case. However, information criteria should not be used, since these typically result in logarithmic rates.*

When accounting for **deterministic components**, the lagged differences are either not affected by a non-zero mean in levels, or can be easily demeaned and deseasonalized, respectively. For the levels y_{t-1}, one must ensure they possess the martingale property, purpose to which we follow Chang (2002, p. 275) and resort to recursive (adaptive) schemes of demeaning, deseasonalizing, or detrending of y_{t-1}.

Recursive adjustment has the additional advantage that it does not matter where deterministics appear - e.g. a linear trend in the data may appear in levels, or due to a non-zero intercept in the cointegrating relation. This is not the case with the Johansen procedure, for instance, where different asymptotic distributions result under different sources of deterministics.

For a non-zero mean, this means that the NIV cointegration test has to be carried out in following test equation:

$$\Delta y_t = \alpha y^{\mu}_{t-1} + \boldsymbol{\beta}'\mathbf{x}^{\mu}_{t-1} + \varepsilon_t, \tag{2.7}$$

where the recursively demeaned lagged level y^{μ}_{t-1} is given for $t \geq 2$ by

$$y^{\mu}_{t-1} = y_{t-1} - \frac{1}{t-1}\sum_{j=1}^{t-1} y_j,$$

and the integrated regressors $\mathbf{x}_{t-1,1}$ may also require demeaning, hence the notation $\mathbf{x}^{\mu}_{t-1}$ in Equation (2.7). Usual projection on a constant is allowed for the integrated regressors, in contrast to the case of the lagged dependent variable. The stationary regressors, being differences, need no adjustment.

For a linear trend, one must use as test equation

$$\Delta y_t = \alpha y^{\tau}_{t-1} + \boldsymbol{\beta}'\mathbf{x}^{\tau}_{t-1} + \varepsilon_t, \tag{2.8}$$

where the recursively detrended lagged level y^{τ}_{t-1} is given for $t \geq 2$ by

$$y^{\tau}_{t-1} = y_{t-1} + \frac{2}{t-1}\sum_{j=1}^{t-1} y_j - \frac{6}{t(t-1)}\sum_{j=1}^{t-1} j\, y_j,$$

and the integrated regressors may be detrended the usual way. The stationary regressors and the regressand Δy_t only require usual demeaning.

For deseasonalizing, note that Kuzin (2005) shows that simply subtracting recursive seasonal means is not admissible, and gives a method that ensures asymptotic equivalence to the case of recursive demeaning. The extension to detrending follows from his work.

Then, one uses as instruments $F(y^{\mu}_{t-1})$ or $F(y^{\tau}_{t-1})$. The employed asymptotic theory holds for these instruments as well, but in terms of recursively demeaned (detrended) Brownian motions, see Chang (2002). Thus, for the case of test Equations (2.7) and (2.8), the results analog to Proposition 2.1, Lemma 2.1 and Proposition 2.2, summarized in the following proposition, can be shown to hold true. Its proof is very similar to that of Proposition 2.2 and not given here.

Proposition 2.3 *Under the assumptions of Lemma 2.1 and recursive demeaning or detrending, it holds for the t statistics from test equations (2.7) or (2.8) as $T \to \infty$:*

a) if $\alpha = 0$, *then*

$$t_{\widehat{\alpha}} \xrightarrow{d} \mathcal{N}(0,1);$$

b) if $\alpha \neq 0$, *then*

$$|t_{\widehat{\alpha}}| \xrightarrow{p} \infty.$$

Proof: *Omitted.*

Remark 2.5 *It can be shown that* $\mathbf{x}_{t-1,1}$ *themselves may be recursively demeaned (detrended) without affecting the asymptotics. This was found to perform better in small samples than projecting* $\mathbf{x}_{t-1,1}$ *on a constant (on a time trend).*

2.3.2 Panel test

We now turn our attention to the panel case and deal with panels containing N vector autoregressive processes representing N cross-sectional units, where N is finite. In what concerns the notation, each used symbol becomes an additional index i, $i = 1, 2, \ldots, N$, representing the respective unit. The innovations $\left(\varepsilon_{i,t}, \boldsymbol{\nu}'_{i,t}\right)'$ are allowed to correlate across units, as specified by following assumption.

Assumption 2.5 *Denote* $\boldsymbol{\epsilon}_t^N = \left(\varepsilon_{1,t}, \boldsymbol{\nu}'_{1,t}, \varepsilon_{2,t}, \boldsymbol{\nu}'_{2,t}, \ldots, \varepsilon_{N,t}, \boldsymbol{\nu}'_{N,t}\right)'$ *and let* $\boldsymbol{\epsilon}_t^N \sim iid(0, \Sigma_N)$, *with* Σ_N *any symmetric, positive definite* $N(K+1) \times N(K+1)$ *matrix. Furthermore, assume* $\exists l > 2$ *such that* $E\left\|\boldsymbol{\epsilon}_t^N\right\|^l < \infty$ *and let* $\boldsymbol{\epsilon}_t^N$ *have absolutely continuous distribution function (w.r.t. Lebesgue measure) such that* $\exists s > 0$ *with* $\phi_N(\boldsymbol{\lambda}) = o\left(\|\boldsymbol{\lambda}\|^s\right)$, *where* ϕ_N *is the corresponding characteristic function.*

Assumptions 2.1, 2.2 and 2.4 are maintained for each single unit. Thus, we do not explicitly allow for cross-unit dynamics, but augmenting the test regressions with lagged differences from other cross-sectional units can obviously be allowed for. Moreover, this augmentation is desirable, since, when ignoring cross-unit dynamics, each unit $\Delta\mathbf{w}_t$ follows (marginally) a general linear process, which requires approximation by means of an autoregressive process of order growing to infinity, see also Remark 2.3.

We assume here the number of cross-sections to be finite. Note that Im and Pesaran (2003) argue that N may grow to infinity, but at a certain maximal rate. However, their argumentation ignores the behavior of the $o_p(1)$ term from Lemma 2.1 as *both* N and T grow to infinity. To avoid this issue, we impose $N < \infty$.

The testing hypotheses are modified as follows:

Null hypothesis: $\alpha_i = 0, \quad i = 1, 2, \ldots, N.$

The null is rejected if equilibrium adjustment is found for at least one unit:

Alternative hypothesis: $\exists\, i,\ 1 \leq i \leq N$ with $\alpha_i \neq 0$.

The panel test statistic is defined as

$$\widetilde{t} = \sum_{i=1}^{N} t^2_{\widehat{\alpha}_i}, \tag{2.9}$$

where the single test statistics $t_{\widehat{\alpha}_i}$ may be computed with recursive demeaning or detrending. Due to Lemma 2.1, one is not constrained to use the same set of explanatory variables for each individual test. This panel test statistic follows a chi-square distribution with N degrees of freedom asymptotically, as stated in the following proposition, and one rejects for too large values of the test statistic. This is because single test statistics are asymptotically independent even if the innovations $\varepsilon_{i,t}$ correlate across units, see the Appendix for proof details. The panel case is where asymptotics of nonlinear transformations of integrated series indeed comes into its own.

Proposition 2.4 *Under the assumptions of Proposition 2.2 with Assumption 2.3 replaced by Assumption 2.5, it holds for $\widetilde{t}$ from (2.9) as $T \to \infty$*

a) if $\alpha_i = 0\, \forall i \in \{1, 2, \ldots, N\}$

$$\widetilde{t} \xrightarrow{d} \chi^2_N,$$

b) if $\exists i$ such that $\alpha_i \neq 0$

$$\widetilde{t} \xrightarrow{p} \infty.$$

Proof: *See the Appendix.*

Remark 2.6 *Proposition 2.4 also allows one to build a simple multi-equation test for single units. Together with Lemma 2.1, this proposition shows that test statistics from different equations* for the same unit *are asymptotically independent as well. Then, one can use as test statistic for a single unit i the sum of the squared t statistics for each equation of unit i. Obviously, a*

χ^2_K distribution results asymptotically. As panel test statistic, one uses the sum of single test statistics for each equation in all units, which follows a χ^2_{NK} distribution.

Remark 2.7 *The same reasoning as in the proof of Proposition 2.4 was used by Chang (2002) in building the panel NIV unit root test. Also, Chang's work, together with Lemma 2.1, shows that panels may be unbalanced in the sense of Assumption 4.1 from Chang (2002).*

Remark 2.8 *The innovations $\boldsymbol{\nu}_{i,t}$ may also correlate across units, with $\boldsymbol{\nu}_{j,t}$ and $\varepsilon_{j,t}$. This is because Lemma 2.1 ensures that the terms containing integrated regressors and lagged differences are asymptotically negligible. Especially useful, the elements of $\mathbf{w}_{i,e,t-1}$ may cointegrate across units.*

Remark 2.9 *Should the dependent variables $y_{i,t}$ cointegrate across units, asymptotic independence is no longer guaranteed. However, Chang and Song (2005) show that independence of single test statistics holds, if the instrument generating functions F_i satisfy certain orthogonality conditions. They suggest the use of Hermite polynomials, but argue that these need rescaling before using them as instrument generating functions (see Chang and Song, 2005, for a complete discussion).*

If knowing that the non-zero α_i are, for instance, negative under the alternative, one-sided testing results in more power. A simple way to build the one-sided panel test is to take the standardized sum of single test statistics,

$$\widetilde{t}^- = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} t_{\widehat{\alpha}_i}. \tag{2.10}$$

Asymptotic standard normality of $\widetilde{t}^-$ is easily proved in the same manner as Proposition 2.4, and one rejects for large negative values of this test statistic.

2.4 Small sample behavior

All simulations are carried out in GAUSS, with 10000 replications for each considered case. Compared to Chang's (2002) simulations, we employ a slightly modified instrument generating function, $F(x) = cx\,e^{-|cx|}$, where c is the inverse of the standard deviation of Δy_t; due to its consistency, it does not affect the asymptotics, in contrast to Chang's (2002) choice of c, which is proportional to $T^{-0.5}$, leading thus to different asymptotics (see also Im and Pesaran, 2003, for a discussion on this subject).

2.4.1 Behavior of the single unit test

First, we study our test for a single unit. We are particularly interested in the effects the lack of exogeneity has on the test statistic. Therefore, we forgo short-run dynamics, but include one lagged difference in the test equation. All integrated variables, y_{t-1} as well as $\mathbf{x}_{t-1,1}$, are recursively demeaned for both partial and complete instrumentalization. For the OLS residual variance estimation, a constant is included in the test equation instead of using recursive demeaning of the integrated variables. Preliminary experiments indicate that better behavior is to be expected if the residual variance estimator is adjusted for degrees of freedom. We take advantage of knowing what the true alternative is, and employ a one-sided test against $\alpha < 0$.

In a first series of experiments, we specify the data generating process with error-correction affecting the other equations, under both the null and the alternative hypothesis. We set $\boldsymbol{\alpha}_e = (-0.1, \ldots, -0.1)$. The covariance matrix of the innovations is set equal to the unity matrix. Under H_1, we let $\alpha = -0.1$ and $\alpha = -0.2$, and, for simplicity, we set the cointegrating vector parameters to $\boldsymbol{\theta} = -\boldsymbol{\alpha}_e$; it can be checked that, with these parameter values, Assumption 2.2 is satisfied for any N and no I(2) process could emerge.

The size and power (at the nominal level of 5%) are given for $K = 1$ and $K = 2$ integrated covariates and sample sizes $T = 100$, $T = 200$ and $T = 500$ in Table 2.1. Both variants of the test are studied: the partially instrumentalized test statistic is denoted by $t^p_{\widehat{\alpha}}$, while the completely instrumentalized one is denoted by $t^c_{\widehat{\alpha}}$.

		$T=100$		$T=200$		$T=500$	
α		$K=1$	$K=2$	$K=1$	$K=2$	$K=1$	$K=2$
0	$t^p_{\widehat{\alpha}}$	6.54	7.44	6.17	6.92	6.52	7.00
	$t^c_{\widehat{\alpha}}$	5.34	6.01	5.15	5.05	4.96	4.64
-0.1	$t^p_{\widehat{\alpha}}$	29.41	28.53	43.41	38.42	70.85	52.46
	$t^c_{\widehat{\alpha}}$	24.01	19.14	31.00	21.39	34.16	18.29
-0.2	$t^p_{\widehat{\alpha}}$	61.49	59.80	86.63	79.57	97.08	88.01
	$t^c_{\widehat{\alpha}}$	48.59	37.29	61.10	41.57	51.96	30.64

Table 2.1: Size and power of single test statistic [%], $\Sigma = I_{K+1}$, $\boldsymbol{\alpha}_e \neq \mathbf{0}$.

We observe the completely instrumentalized test to be mildly oversized for $T = 100$. For larger sample sizes, the size is close to the nominal level. For all studied sample sizes, the size distortions are of larger magnitude for the partially instrumentalized test and increase with growing K. On the other hand, the partially instrumentalized test is clearly superior in terms of power: $t^p_{\widehat{\alpha}}$ dominates $t^c_{\widehat{\alpha}}$ for all sample sizes and all values of K, being up to four times more powerful. The power of both $t^p_{\widehat{\alpha}}$ and $t^c_{\widehat{\alpha}}$ decreases with increasing K, mirroring the situation under the null. As expected, the power increases with growing sample size. The test by Banerjee, Dolado and Mestre (1998) is more powerful (even though exogeneity is not provided for, its rejection frequency, for instance for $T = 500$ and $\alpha = -0.1$, is 100%), but this should not come as a surprise, since using instrumental estimation instead of OLS usually leads to less power; in our particular situation, the NIV test statistic also diverges slower under the alternative than the one based on OLS.

For a second series of experiments, $\mathbf{x}_{t-1,1}$ is not affected by error correction, i.e. we set $\boldsymbol{\alpha}_e = \mathbf{0}$. In exchange, we let the lack of exogeneity to be caused by correlation between innovations ε_t and $\boldsymbol{\nu}_t$. More precisely, we use a covariance matrix of the innovations with constant correlation, i.e. we set Σ from Assumption 2.3 to equal $\Sigma_{(K+1)\times(K+1)} = \{\rho_{ij}\}_{1\leq i,j\leq K+1}$, with $\rho_{ij} = 1$ if $i = j$, and $\rho_{ij} = \rho$ if $i \neq j$. For Σ to be positive definite, it must hold that $-1/K < \rho < 1$. We choose for ρ the value of 0.5, to account for positive correlation often observed between macroeconomic time series. The rejection frequencies for the 5% level are given in Table 2.2.

		$T = 100$		$T = 200$		$T = 500$	
α		$K = 1$	$K = 2$	$K = 1$	$K = 2$	$K = 1$	$K = 2$
0	$t^p_{\widehat{\alpha}}$	6.07	6.84	5.30	6.25	4.86	5.68
	$t^c_{\widehat{\alpha}}$	6.42	6.52	5.35	5.39	5.07	5.01
-0.1	$t^p_{\widehat{\alpha}}$	22.50	19.96	38.26	30.00	69.54	48.88
	$t^c_{\widehat{\alpha}}$	19.72	16.06	28.96	18.81	35.81	18.56
-0.2	$t^p_{\widehat{\alpha}}$	51.07	41.17	79.09	61.78	96.06	77.21
	$t^c_{\widehat{\alpha}}$	40.37	27.58	50.36	30.25	46.15	23.33

Table 2.2: Size and power of single test statistic [%], $\rho = 0.5$, $\boldsymbol{\alpha}_e = \mathbf{0}$.

Here, the size distortions of the completely instrumentalized test are larger, especially for $T = 100$, but decrease with growing T. The overall image is basically the same, although the size distortions of the partially instrumentalized test decrease faster for increasing sample size than before. The power is marginally lower for each studied case.

Finally, both sources of lack of exogeneity are combined and the respective results are given in Table 2.3. The findings of the first two series of experiments are confirmed: the power is much higher for $t^p_{\widehat{\alpha}}$ than for $t^c_{\widehat{\alpha}}$, while $t^c_{\widehat{\alpha}}$ has better size properties. The power is again marginally lower than for each source of lack of exogeneity studied separately.

α		$T = 100$		$T = 200$		$T = 500$	
		$K = 1$	$K = 2$	$K = 1$	$K = 2$	$K = 1$	$K = 2$
$\alpha = 0$	$t^p_{\widehat{\alpha}}$	5.99	6.98	6.02	6.61	5.99	7.11
	$t^c_{\widehat{\alpha}}$	5.82	6.94	5.25	5.99	5.07	5.53
-0.1	$t^p_{\widehat{\alpha}}$	20.76	17.23	31.43	21.51	51.55	28.58
	$t^c_{\widehat{\alpha}}$	16.98	13.98	21.74	13.65	19.30	11.58
-0.2	$t^p_{\widehat{\alpha}}$	46.57	35.51	70.38	47.60	89.01	57.84
	$t^c_{\widehat{\alpha}}$	35.04	22.78	39.26	22.34	32.15	16.67

Table 2.3: Size and power of single test statistic [%], $\rho = 0.5$, $\boldsymbol{\alpha_e} \neq \mathbf{0}$.

Summing up, the partially instrumentalized test statistic performs well in terms of power, and is slightly oversized under the null hypothesis, although, compared to tests that do not account for cross-dependence, these size distortions are not worth mentioning. Opposed to that, the completely instrumentalized test statistic does not behave satisfactorily under the alternative hypothesis; in fact, the power is so low that the better size properties are overshadowed. Both variants of the test statistic are robust to absence of exogeneity.

2.4.2 Behavior of the panel test

For the panel situation, we examine how well the asymptotic independence of single test statistics is preserved in small samples. We first study a panel of $N = 2$ units in detail, and turn our attention to larger cross-sectional dimension afterwards. Having generated a one-sided alternative for the Monte Carlo analysis, we use the test statistic $\widetilde{t}^-$ from Equation (2.10) to obtain more power.

For this series of experiments, we use the same form of covariance matrix as above; it allows in the panel situation for lack of exogeneity in each unit as well as for dependence across units; $\Sigma_N = Cov\left(\varepsilon_{1,t}, \boldsymbol{\nu}'_{1,t}, \varepsilon_{2,t}, \boldsymbol{\nu}'_{2,t}\right)'$ is a

$(2K+2) \times (2K+2)$ matrix,

$$Cov\left(\varepsilon_{1,t}, \boldsymbol{\nu}_{1,t}', \varepsilon_{2,t}, \boldsymbol{\nu}_{2,t}'\right)' = \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho & \rho \\ \rho & 1 & \rho & \cdots & \rho & \rho \\ \rho & \rho & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \rho & \rho \\ \rho & \rho & \cdots & \rho & 1 & \rho \\ \rho & \rho & \cdots & \rho & \rho & 1 \end{pmatrix}.$$

Of course, any positive definite matrix Σ_N is allowed by our results. However, aside from an economic justification (see O'Connell, 1998), this is a particularly persistent correlation structure (within as well as across units): the largest eigenvalue of Σ is $O(NK)$. In contrast, Chang's (2002) Monte Carlo setup for the unit root test uses random eigenvalues between 0.1 and 1, which leads to weak cross-correlation, see also Im and Pesaran (2003).

The correlation ρ is again chosen to be 0.5; this implies a correlation of 0.5 between the innovations of the two units.

We only report the results in the situation more relevant for practical applications, i.e. with both sources of endogeneity. The size and power of the panel test are given in Table 2.4 for $N = 2$.

		$T=100$		$T=200$		$T=500$	
α		$K=1$	$K=2$	$K=1$	$K=2$	$K=1$	$K=2$
$\alpha=0$	$t^p_{\widehat{\alpha}}$	5.83	6.99	6.01	6.85	6.22	6.94
	$t^c_{\widehat{\alpha}}$	5.10	6.43	5.10	5.89	5.31	5.97
-0.1	$t^p_{\widehat{\alpha}}$	32.52	24.86	49.63	32.77	73.94	40.64
	$t^c_{\widehat{\alpha}}$	25.48	18.30	30.52	18.73	26.58	13.76
-0.2	$t^p_{\widehat{\alpha}}$	69.68	54.79	91.38	68.91	97.40	74.48
	$t^c_{\widehat{\alpha}}$	51.09	32.86	55.58	30.55	43.82	21.07

Table 2.4: Size and power of panel test, $\rho = 0.5$, $\boldsymbol{\alpha_e} \neq \mathbf{0}$, $N = 2$.

The power increases compared to the single unit case, even for the partially instrumentalized test statistic. The most interesting fact, however, is that the magnitude of the size distortions under the null hypothesis is approximatively the same as for the single test, in spite of a correlation of 0.5 between the innovations of the two units! This suggests that the property of asymptotic independence of single test statistics is well maintained in small samples. This intuition is confirmed by Table 2.5, where we report mean and standard deviation of a single test statistic (μ_i and σ_i), as well as mean and standard deviation of the panel test statistic (μ and σ) for $N = 2$, together with the correlation (ϱ) between the two single test statistics based on which the panel test statistic is computed (Monte Carlo estimates based on 10000 samples).

			$T_i = 100$		$T_i = 200$		$T_i = 500$	
N			$K_i = 1$	$K_i = 2$	$K_i = 1$	$K_i = 2$	$K_i = 1$	$K_i = 2$
1	$t^p_{\hat\alpha}$	μ_i	0.055	0.016	0.029	0.001	−0.016	−0.056
		σ_i	1.098	1.131	1.072	1.082	1.053	1.057
	$t^c_{\hat\alpha}$	μ_i	0.065	0.001	0.041	0.000	−0.005	−0.027
		σ_i	1.093	1.101	1.048	1.058	1.018	1.029
2	$t^p_{\hat\alpha}$	μ	0.078	0.022	0.041	0.002	−0.023	−0.079
		σ	1.109	1.144	1.073	1.094	1.049	1.065
		ϱ	0.020	0.022	0.014	0.021	−0.007	0.015
	$t^c_{\hat\alpha}$	μ	0.092	0.002	0.058	0.000	−0.007	−0.038
		σ	1.106	1.105	1.063	1.067	1.024	1.036
		ϱ	0.023	0.006	0.029	0.016	0.012	0.014

Table 2.5: Characteristics of single and panel ($N = 2$) test statistics under the null hypothesis.

Indeed, the correlation between single test statistics is practically negligible. Even for the worst case ($T = 100$ and $K = 1$), it is rather small: 0.023 compared to its cause, $\rho = 0.5$, and decreases with growing T. The moments of the asymptotic distribution are better approximated with complete instrumentalization than with partial instrumentalization, thus explaining the better size properties of $t^c_{\hat\alpha}$.

For the case N > 2, we note that the main problem resides not with cross-dependence, but with the distortions single test statistics exhibit. Non-zero mean of single test statistics has the worst effect, since, due to the definition of the panel test statistic, it is multiplied by $\sqrt{N}$. Inflated standard deviations, however, do not affect that much the size of the panel test statistic as the number of units grows. A small-sample correction for single test statistics seems appropriate, especially for centering. Finally, for very large N, even small correlations between single test statistics will have negative effect on the size of the panel test, see also Im and Pesaran (2003).

Table 2.6 studies the size properties of the panel test for $N = 5$ and $N = 10$, together with the power under the alternatives $\alpha_i = -0.1$ and $\alpha_i = -0.2$. The results suggest that complete instrumentalization should be used, if at all, in panels with larger cross-sectional dimension, where a large N may compensate for the lack of power of single tests, and the better size properties of single tests result in a better behavior of the panel test under the null hypothesis.

			$T_i = 100$		$T_i = 200$		$T_i = 500$	
N	α_i		$K_i = 1$	$K_i = 2$	$K_i = 1$	$K_i = 2$	$K_i = 1$	$K_i = 2$
5	0	$t^p_{\hat{\alpha}}$	6.36	7.76	6.27	6.98	6.87	8.25
		$t^c_{\hat{\alpha}}$	5.33	6.95	4.85	5.95	5.49	6.28
	−0.1	$t^p_{\hat{\alpha}}$	55.31	42.92	78.52	56.18	93.70	65.59
		$t^c_{\hat{\alpha}}$	42.48	28.82	51.41	29.25	43.11	21.55
	−0.2	$t^p_{\hat{\alpha}}$	92.84	80.81	99.36	89.34	99.47	90.81
		$t^c_{\hat{\alpha}}$	75.94	51.64	77.87	49.58	65.65	34.40
10	0	$t^p_{\hat{\alpha}}$	7.44	8.93	6.67	7.44	7.06	9.52
		$t^c_{\hat{\alpha}}$	5.05	6.67	5.18	6.10	6.33	6.56
	−0.1	$t^p_{\hat{\alpha}}$	74.15	60.11	91.73	73.59	97.83	81.12
		$t^c_{\hat{\alpha}}$	60.08	40.98	68.12	42.51	59.05	32.04
	−0.2	$t^p_{\hat{\alpha}}$	98.06	91.84	99.91	95.38	99.86	95.42
		$t^c_{\hat{\alpha}}$	89.43	69.62	89.32	66.17	80.77	49.94

Table 2.6: Size and power of panel test, $N = 5$ and $N = 10$.

For $N = 10$, the power, although increasing compared to $N = 2$ and $N = 5$, is still relatively low even for $T = 500$. For partial instrumentalization, the size distortions are caused by the mean of the test statistic. For instance, with $N = 10$, $T = 100$ and $K = 1$, this mean equals 0.162 for the partially instrumentalized test (estimation from the same Monte Carlo experiment), larger than the mean of the single test statistic by a proportionality factor of about 2.945, close to $\sqrt{10}$. The size properties become less acceptable with growing K and N. On the other hand, the power increases with N, as expected. **To sum up**, partial instrumentalization is to be preferred as long as $N \ll T$.

2.5 Conclusions

Until recently, panel cointegration studies have been carried out under the assumption of independent units. This not being a plausible assumption, especially in macroeconomic panels, it is expected that such studies lead to biased conclusions. The present paper proposes a test that does not exhibit such shortcoming.

Our test is obtained from the error correction representation of Banerjee, Dolado and Mestre (1998), to which the nonlinear instrumental variable method was applied in a manner similar to Chang (2002).

The proposed test statistic is shown to be asymptotically standard normal distributed and requires no exogeneity assumptions. We find, however, that the residual variance should not be estimated using NIV estimators of the parameters, since these may not converge fast enough to ensure consistent residuals. Instead, we use the OLS residual variance estimator.

In cross-correlated as well as in cross-cointegrated panels, individual test statistics are shown to be asymptotically independent; thus, a panel cointegration test robust to cross-dependence can be built. Up to a certain degree, panels may also be unbalanced, and no N asymptotics is required.

The included Monte Carlo evidence shows that, in small samples, the property of asymptotic independence of single test statistics is well maintained. This is true only to a lesser degree for the approximation of the small sample distribution of single test statistics by their asymptotic distribution. Our test should be used in panels with small cross-sectional dimension (relative to the time dimension); a typical application would consist in multi-country studies, with several countries observed over a couple of decades.

Appendix

The proofs of the results stated in the paper require following lemma.

Lemma A Under the assumptions of Proposition 2.1, it holds as $T \to \infty$:

A.1

$$\frac{1}{T^{0.5}} \sum_{t=1}^{T} F^2 (y_{t-1}) \xrightarrow{d} \mathcal{L}(1,0) \int_{-\infty}^{\infty} F^2(s)\, ds\,;$$

A.2

$$\frac{1}{T^{0.5}} \sum_{t=1}^{T} y_{t-1} F(y_{t-1}) \xrightarrow{d} \mathcal{L}(1,0) \int_{-\infty}^{\infty} sF(s)\, ds\,;$$

A.3

$$\frac{1}{T^{0.25}} \sum_{t=1}^{T} F(y_{t-1})\, \varepsilon_t \xrightarrow{d} \sigma_\varepsilon \sqrt{\mathcal{L}(1,0) \int_{-\infty}^{\infty} F^2(s)\, ds} \cdot W(1)\;;$$

A.4

$$\frac{1}{T^{0.5}} \sum F(y_{t-1})\, \mathbf{x}'_{t-1,0} = O_p(1)\;,$$

where $W(r)$ is a standard Brownian motion and $\mathcal{L}(t,s)$ is the local Brownian time associated with another Brownian motion U, independent of W and having as variance the long-run variance of Δy_{t-1}:

$$\mathcal{L}(t,s) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1\left(|U(r) - s| < \epsilon\right) dr,$$

with $1(\cdot)$ the usual indicator function, $1(A) = 1$, if proposition A is true, and 0, otherwise.
If, additionally, $\boldsymbol{\alpha}_e = \mathbf{0}$, it holds

A.5
$$\frac{1}{T}\sum_{t=1}^{T} F\left(y_{t-1}\right)\mathbf{x}'_{t-1,1} = O_p\left(1\right);$$

if $\boldsymbol{\alpha}_e \neq \mathbf{0}$, a different behavior of this sample cross-moment can emerge, depending on $\boldsymbol{\theta}$:

A.6
$$\frac{1}{T}\sum_{t=1}^{T} F\left(y_{t-1}\right)\mathbf{x}'_{t-1,1} = O_p\left(T^{-0.5}\right).$$

Proof of Lemma A

A.1 follows from de Jong and Wang (2005), Theorem 2, since F^2 satisfies their conditions when F obeys our Assumption 2.4.

A.2 also follows from de Jong and Wang (2005).

A.3 is proven by Demetrescu (2006), see his Lemma 1, item d).

A.4 is a direct consequence of Demetrescu's Lemma 1, item e).

A.5 is shown to hold true under no cointegration (i.e. $\alpha = 0$ and $\boldsymbol{\alpha}_e = \mathbf{0}$) by Chang, Park and Phillips (2001), Lemma 5.

A.6 Under cointegration of y_{t-1} and $\mathbf{x}_{t-1,1}$ (as implied by $\boldsymbol{\alpha}_e \neq \mathbf{0}$), each element of $\mathbf{x}_{t-1,1}$ can be expressed as a linear combination of I(1) variables that are either cointegrated with y_{t-1} or not. When at least one of the I(1) variables that are not cointegrated with y_{t-1} is present in the linear combinations, the cross-moment has order $O_p(1)$ due to A.5; otherwise, when elements of $\mathbf{x}_{t-1,1}$ equal y_{t-1} plus I(0) noise, the $O_p(T^{-0.5})$ order emerges due to A.2 and A.4.

Proof of Proposition 2.1

a) Recall, $\widehat{\alpha} - \alpha = Q^{-1}M$, with M and Q defined in the text. Let us now examine the behavior of M. The second term on the right-hand side of the equation defining M can be written as

$$\left(\sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}_{t-1}'\right) D_T^{-1} D_T \left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right)^{-1} D_T D_T^{-1} \left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\varepsilon_t\right)$$

with D_T a $((K+1)p+K) \times ((K+1)p+K)$ diagonal matrix partitioned according to the stationary and integrated components of $\mathbf{x}_{t-1}$:

$$D_T = \begin{pmatrix} T^{0.5} & 0 \\ 0 & T \end{pmatrix}, \quad D_T^{-1} = \begin{pmatrix} T^{-0.5} & 0 \\ 0 & T^{-1} \end{pmatrix}.$$

It follows that

$$\begin{aligned} &\left(\sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}_{t-1}'\right) \begin{pmatrix} T^{-0.5} & 0 \\ 0 & T^{-1} \end{pmatrix} \\ &= \left(T^{-0.5} \sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}_{t-1,0}', T^{-1} \sum_{t=1}^{T} F\left(y_{t-1}\right) \mathbf{x}_{t-1,1}'\right) \\ &= \left(O_p\left(1\right), O_p\left(1\right)\right) \end{aligned}$$

and

$$\begin{pmatrix} T^{-0.5} & 0 \\ 0 & T^{-1} \end{pmatrix} \left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\varepsilon_t\right) = \begin{pmatrix} T^{-0.5} \sum_{t=1}^{T} \mathbf{x}_{t-1,0}\varepsilon_t \\ T^{-1} \sum_{t=1}^{T} \mathbf{x}_{t-1,1}\varepsilon_t \end{pmatrix} = \begin{pmatrix} O_p\left(1\right) \\ O_p\left(1\right) \end{pmatrix},$$

see Lemma A.4 and A.5. Then,

$$D_T \left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right)^{-1} D_T = \left(D_T^{-1} \left(\sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{x}_{t-1}'\right) D_T^{-1}\right)^{-1}$$

which equals

$$\left(\left(\begin{array}{cc} T^{-0.5} & 0 \\ 0 & T^{-1} \end{array}\right)\left(\begin{array}{cc} \sum \mathbf{x}_{t-1,0}\mathbf{x}_{t-1,0}' & \sum \mathbf{x}_{t-1,0}\mathbf{x}_{t-1,1}' \\ \sum \mathbf{x}_{t-1,1}\mathbf{x}_{t-1,0}' & \sum \mathbf{x}_{t-1,1}\mathbf{x}_{t-1,1}' \end{array}\right)\left(\begin{array}{cc} T^{-0.5} & 0 \\ 0 & T^{-1} \end{array}\right)\right)^{-1}$$

$$= \left(\begin{array}{cc} T^{-1}\sum \mathbf{x}_{t-1,0}\mathbf{x}_{t-1,0}' & T^{-1.5}\sum \mathbf{x}_{t-1,0}\mathbf{x}_{t-1,1}' \\ T^{-1.5}\sum \mathbf{x}_{t-1,1}\mathbf{x}_{t-1,0}' & T^{-2}\sum \mathbf{x}_{t-1,1}\mathbf{x}_{t-1,1}' \end{array}\right)^{-1}$$

$$= \left(\begin{array}{cc} O_p(1) & O_p(T^{-0.5}) \\ O_p(T^{-0.5}) & O_p(1) \end{array}\right)^{-1} = \left(\begin{array}{cc} O_p(1) & O_p(T^{-0.5}) \\ O_p(T^{-0.5}) & O_p(1) \end{array}\right),$$

due to continuity of matrix inversion and nonsingularity. Hence,

$$\begin{aligned} M &= \sum_{t=1}^{T} F(y_{t-1})\varepsilon_t - (O_p(1), O_p(1))\left(\begin{array}{cc} O_p(1) & O_p(T^{-0.5}) \\ O_p(T^{-0.5}) & O_p(1) \end{array}\right)\left(\begin{array}{c} O_p(1) \\ O_p(1) \end{array}\right) \\ &= \sum_{t=1}^{T} F(y_{t-1})\varepsilon_t - O_p(1) = O_p\left(T^{0.25}\right). \end{aligned}$$

For Q, we only need to examine

$$\left(\begin{array}{cc} T^{-0.5} & 0 \\ 0 & T^{-1} \end{array}\right)\left(\sum_{t=1}^{T} \mathbf{x}_{t-1}y_{t-1}\right) = \left(\begin{array}{c} T^{0.5}\left(T^{-1}\sum_{t=1}^{T} \mathbf{x}_{t-1,0}y_{t-1}\right) \\ T\left(T^{-2}\sum_{t=1}^{T} \mathbf{x}_{t-1,1}y_{t-1}\right) \end{array}\right) = \left(\begin{array}{c} O_p(T^{0.5}) \\ O_p(T) \end{array}\right),$$

which, under no cointegration, leads to

$$\begin{aligned} Q &= \sum_{t=1}^{T} F(y_{t-1})y_{t-1} - (O_p(1), O_p(1))\left(\begin{array}{cc} O_p(1) & O_p(T^{-0.5}) \\ O_p(T^{-0.5}) & O_p(1) \end{array}\right)\left(\begin{array}{c} O_p(T^{0.5}) \\ O_p(T) \end{array}\right) \\ &= O_p\left(T^{0.5}\right) - (O_p(1), O_p(1))\left(\begin{array}{c} O_p(T^{0.5}) \\ O_p(T) \end{array}\right) \\ &= O_p\left(T^{0.5}\right) - (O_p(1) + O_p(T)) = O_p(T), \end{aligned}$$

from which the convergence rate for $\widehat{\alpha}$ follows directly,

$$\widehat{\alpha} - \alpha = O_p\left(T^{-1}\right)O_p\left(T^{0.25}\right) = O_p\left(T^{-0.75}\right).$$

For $\boldsymbol{\beta}$, we have

$$\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} = J^{-1}R$$

with R a column vector:

$$R = \sum_{t=1}^{T} \mathbf{x}_{t-1}\varepsilon_t - \sum_{t=1}^{T} \mathbf{x}_{t-1}y_{t-1}\left(\sum_{t=1}^{T} F\left(y_{t-1}\right)y_{t-1}\right)^{-1}\sum_{t=1}^{T} F\left(y_{t-1}\right)\varepsilon_t$$

and J a matrix

$$J = \sum_{t=1}^{T} \mathbf{x}_{t-1}\mathbf{x}_{t-1}' - \sum_{t=1}^{T} \mathbf{x}_{t-1}y_{t-1}\left(\sum_{t=1}^{T} F\left(y_{t-1}\right)y_{t-1}\right)^{-1}\sum_{t=1}^{T} F\left(y_{t-1}\right)\mathbf{x}_{t-1}'.$$

Split R and J corresponding to the stationary and integrated regressors. Then, it is straightforward to check that for $R = (R_0', R_1')'$ it holds

$$R_0 = O_p\left(T^{0.75}\right) \text{ and } R_1 = O_p\left(T^{1.75}\right).$$

Then, for

$$J = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

it holds due to Lemma A that $A = O_p(T)$, $B = O_p(T^{1.5})$, $C = O_p(T^2)$ and $D = O_p(T^{2.5})$. Using formulae for inverting partitioned matrices (e.g. Lütkepohl, 1996, p. 147), one obtains, after some algebra,

$$J^{-1} = \begin{pmatrix} O_p(T^{-1}) & O_p(T^{-2}) \\ O_p(T^{-1.5}) & O_p(T^{-2.5}) \end{pmatrix},$$

from which the desired convergence rates follow.

b) The result follows along the same lines, but now A.6 could hold instead of A.5 (see the proof of Lemma A for details). While M remains in any case of order $O_p(T^{0.25})$, Q can be of order $O_p(T^{0.5})$ instead of $O_p(T)$, resulting in a convergence order for $\widehat{\alpha}$ of $O_p(T^{0.25})$. In what concerns $\widehat{\boldsymbol{\beta}}$, B may be $O_p(T)$ and D may be $O_p(T^2)$, if A.6 holds. Thus, a behavior similar to that of $\widehat{\alpha}$ emerges for $\widehat{\boldsymbol{\beta}}_1$, while the behavior of $\widehat{\boldsymbol{\beta}}_0$ is unaffected.

Proof of Lemma 2.1 From the proof of Proposition 2.1, it follows

$$M = \frac{1}{T^{0.25}}\sum_{t=1}^{T} F\left(y_{t-1}\right)\varepsilon_t - o_p\left(1\right).$$

We also have by arguments similar to those in the proof of Proposition 2.1 that

$$P=\sum_{t=1}^{T}F\left(y_{t-1}\right)^{2}-\left(O_{p}\left(1\right),O_{p}\left(1\right)\right)\left(\begin{array}{cc}O_{p}\left(1\right) & O_{p}\left(T^{-0.5}\right)\\ O_{p}\left(T^{-0.5}\right) & O_{p}\left(1\right)\end{array}\right)\left(\begin{array}{c}O_{p}\left(1\right)\\ O_{p}\left(1\right)\end{array}\right),$$

and thus

$$\frac{1}{T^{0.5}}P=\frac{1}{T^{0.5}}\sum_{t=1}^{T}F\left(y_{t-1}\right)^{2}-o_{p}\left(1\right),$$

which leads to the desired result, the numerator of the t statistic being different from zero with probability 1. It is obvious that this relation holds of irrespective of which behavior, A.5 or A.6, holds.

Proof of Proposition 2.2 The result follows directly from Lemma 2.1, if joint convergence of A.3 and A.1 holds. This is indeed the case, since the proof of A.3 given by Demetrescu (2006) establishes A.3 as an implication of A.1.

Proof of Proposition 2.4 Chang (2002) argues that, as $T\to\infty$,

$$\frac{1}{\sqrt[4]{T}}\sum_{t=1}^{T}F_{i}\left(y_{i,t-1}\right)\varepsilon_{i,t}\approx_{d}\sqrt[4]{T}\int_{0}^{1}F_{i}\left(\sqrt{T}U_{i}\right)\,dW_{i},$$

where $\approx_d$ stands for equivalence in distribution. This also applies for recursively demeaned (detrended) data, y_{t-1}^{μ} or y_{t-1}^{τ}, but in relation with the corresponding recursively demeaned (detrended) Brownian motions. The left-hand side of equation above is the numerator of the individual test statistic. For two units $i\neq j$, the equivalent right-hand side representations are independent if and only if their quadratic covariation, given by

$$\sigma_{ij}\sqrt{T}\int_{0}^{1}F_{i}\left(\sqrt{T}U_{i}\right)F_{j}\left(\sqrt{T}U_{j}\right)\,ds,$$

where σ_{ij} is the covariance of $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$, disappears almost surely as $T\to\infty$ (see Chang, Park and Phillips, 2001). For independent units, this condition holds trivially. But even if $\sigma_{ij}\neq 0$, it is known (see Kasahara and Kotani,

1979) that

$$\int_0^1 F_i\left(\sqrt{T}U_i\right) F_j\left(\sqrt{T}U_j\right) ds = O_p\left(\frac{\ln T}{T}\right).$$

The denominators are asymptotically uncorrelated, since, as Chang, Park and Phillips (2001, Lemma 5, item j) show, $T^{-0.5}\sum_{t=1}^{T} F_i^2\left(y_{i,t-1}\right) F_j^2\left(y_{j,t-1}\right) = o_p(1)$, the function $F^2(\cdot)$ being itself regularly integrable, if $F(\cdot)$ is. This establishes asymptotic independence, from which the first result follows. Consistency is obvious given Proposition 2.1, item b).

Chapter 3

Combining Significance of Correlated Statistics with Application to Panel Data

3.1 Introduction

In clinical studies, there is a long tradition of combining p-values arising from test statistics of different experiments. In a meta analysis, the outcomes of several experiments are used to test for an overall effect. In particular, in case of very different test statistics, a direct combination of them may be impossible, and one has to rely on the combination of their respective p-values. The idea to do so is attributed to Fisher - see, for instance, Fisher (1954, p. 100).

Recently, the combination of significance levels has been advocated in econometrics, too. A typical application would be to build a joint significance test for parameter values across units. For instance, Maddala and Wu (1999) and Choi (2001) propose the combination of p-values computed from individual cross-sections in the context of panel unit root testing[1]. This allows for the combination of arbitrary tests, can be used with unbalanced panels and resorts to standard limiting distributions (χ^2 or $N(0,1)$) that do not require a large amount of tables. The asymptotics assume growing time dimension and a finite number of cross-sections (large T, small N), which is typical for many multi-country studies. Still, the case $N \to \infty$ can be incorporated, see Choi (2001). The major drawback is that asymptotic theory typically builds on the assumption of independent test statistics. Clearly, this is not a valid assumption working with macroeconomic panels. To allow for cross-sectional

[1] For recent surveys on panel unit root tests see Banerjee (1999) and Baltagi and Kao (2000).

dependence when testing for unit roots, Maddala and Wu (1999) suggest the use of panel bootstrap. Not wishing to resort to resampling methods, Chang (2002) suggests nonlinear instrument estimation of the Dickey-Fuller test regression resulting in asymptotically independent test statistics, while Bai and Ng (2004) focus on the source of cross-dependence[2] by means of a factor model.

Being concerned with outlier detection, Carrion-i-Silvestre (2003) adopts a different approach of combining significance levels under dependence. This approach is due to Hartung (1999) who modifies the so-called inverse normal method[3] in order to robustify against correlation. The inverse normal method works with transformations of p-values, say t_i, which follow a standard normal distribution by construction and are computed for N statistics, $i = 1, \ldots, N$. Hartung (1999, p. 850) then exploits the fact that, under normality, "dependency in the original test statistics (...) is equivalent to some correlation“, and, by assuming a certain structure (see Section 3.2) of the covariance matrix of the transformed p-values, t_i, he shows how to correct for cross-correlation.

The present work tackles two aspects of this modification of the inverse normal method. On the one hand, the assumed correlation structure is likely to be violated in applied work. On the other hand, the marginal normal distribution of t_i does not imply multivariate normality of $t_1, \ldots, t_N$. Hence, a linear combination of the probits does not necessarily follow a normal distribution as required for the modified inverse normal method. Our contribution is, correspondingly, twofold. We first show the modified inverse normal method to work under a wide class of covariance matrices of the probits. Then, in terms of copulas, we provide a necessary and sufficient condition for the theoretical justification of Hartung's suggestion. Although this condition does not hold, for instance, in case of Dickey-Fuller unit root tests, experimental evidence shows that correcting for cross-correlation leads in finite samples to significant improvements over the naive inverse normal method. Hassler and Tarcolea (see Chapter 4) recently applied this method to OECD long-term interest rates, finding only rather weak evidence in favor of stationarity when allowing for cross-correlation.

This chapter is structured as follows. Section 3.2 reviews the modified inverse normal method and discusses conditions for its validity. Then, we focus on panel tests as an application of this method, and give Monte Carlo examinations of our theoretical results. The main findings are summarized in

[2] Their approach allows for a wider class of inference than only unit root testing.

[3] The method was introduced in case of independent statistics by Liptak (1958); further refinements were given, among others, by Van Zwet and Oosterhoff (1967).

Section 3.4.

3.2 The modified inverse normal method

Before presenting and discussing the modification of the inverse normal method proposed by Hartung (1999), we introduce some notation.

Let S_i $(i = 1, ..., N)$ stand for statistics testing the null hypotheses

$$H_{i0}: \ \theta_i = \theta_{i0}\,, \quad \theta_i \in \mathbb{R}\,, \quad i = 1, \ldots, N\,,$$

against the one-sided alternatives

$$H_{i1}: \theta_i < \theta_{i0}.$$

It is assumed that S_i has a continuous distribution function, F_{i0}, under H_{i0}. The corresponding p-values p_i are used to define the so-called probits t_i,

$$t_i = \Phi^{-1}(p_i) \ \text{ with } p_i = F_{i0}(S_i), \tag{3.1}$$

where Φ^{-1} is the inverse of the standard normal cumulative distribution function (cdf). Each probit follows a standard normal distribution by definition, if the null hypothesis holds.

The (weighted) inverse normal method consists of linearly combining the probits; for independent units, one obtains a normally distributed combined test statistic. Hartung (1999) extends the usual weighted inverse normal method in order to account for dependent probits. Dependency in the original statistics S_i leads to dependency in the corresponding probits t_i $(i = 1, ..., N)$. It is assumed that this can be expressed as correlation. In this case, the covariance equals the correlation, because probits are, by construction, standardized, so one has $Cov(t_i, t_j) = \rho_{ij}$. More precisely, constant correlation is assumed,

$$Cov(t_i, t_j) = \rho, \qquad \text{for } i \neq j, \quad i, j = 1, ..., N, \tag{3.2}$$

with $-\frac{1}{N-1} < \rho < 1$, condition that guarantees positive definiteness of the covariance matrix following from (3.2).

If ρ was known, one might compute with weights $\lambda_i \in \mathbb{R}$, $\sum_{i=1}^{N} \lambda_i \neq 0$,

$$t(\rho) = \frac{\sum_{i=1}^{N} \lambda_i t_i}{\sqrt{(1-\rho) \sum_{i=1}^{N} \lambda_i^2 + \rho \left(\sum_{i=1}^{N} \lambda_i \right)^2}} . \tag{3.3}$$

For $\rho = 0$ and $\lambda_i = 1$ this reproduces the test statistic favoured by Choi (2001). Under the joint null hypothesis,

$$H_0 : \ \theta_i = \theta_{i0} \text{ for all } i = 1, 2, \ldots, N ,$$

the combined statistic $t(\rho)$ follows a standard normal distribution if the probits are from a multivariate normal distribution, and the joint test rejects for large negative values of the test statistic.

In practice, ρ is not zero and not known but has to be estimated on the basis of the observed t_i. Plugging in a consistent (as $N \to \infty$) estimator $\widehat{\rho}^*$, Hartung (1999) suggests

$$t(\widehat{\rho}^*, \kappa) = \frac{\sum_{i=1}^{N} \lambda_i t_i}{\sqrt{\sum_{i=1}^{N} \lambda_i^2 + \left[\left(\sum_{i=1}^{N} \lambda_i \right)^2 - \sum_{i=1}^{N} \lambda_i^2 \right] \left[\widehat{\rho}^* + \kappa \sqrt{\frac{2}{N+1}} \, (1 - \widehat{\rho}^*) \right]}} , \tag{3.4}$$

where

$$\widehat{\rho}^* = \max \left(-\frac{1}{N-1}, \widehat{\rho} \right) \quad \text{with } \widehat{\rho} = 1 - \frac{1}{N-1} \sum_{i=1}^{N} \left(t_i - \frac{1}{N} \sum_{i=1}^{N} t_i \right)^2 ,$$

and $\kappa > 0$ is a parameter regulating in small samples the actual significance level, chosen in Hartung's simulations as $\kappa = \kappa_1 = 0.2$ or as $\kappa = \kappa_2 = 0.1 \cdot (1 + \frac{1}{N-1} - \widehat{\rho}^*)$. This estimator is justified by the properties of quadratic forms of a multivariate normal distribution possessing the assumed covariance matrix. For further construction details and comments see Hartung (1999). Again, given a consistent $\widehat{\rho}^*$ and a multivariate normal distribution of $t_1, \ldots, t_N$, the combined test statistic $t(\widehat{\rho}^*, \kappa)$ results in an approximate standard normal distribution under the null hypothesis.

In practical applications, it may well happen that the data generating process does not lead to constant correlation of the probits. Hartung provides in his paper simulations which show that a certain degree of deviation from the constant-correlation assumption is tolerable. We go a step further and, in the following Proposition, provide technical conditions under which the modified inverse normal method is robust to deviations from the constant correlation assumption.

Proposition 3.1 *Let* $\Sigma = \{\rho_{ij}\}_{i,j=1,\dots,N}$ *be the covariance matrix of the probits* $\mathbf{t} = \{t_i\}_{i=1,\dots,N}$ *and assume*

$$\lim_{N\to\infty} \frac{1}{N(N-1)} \sum \sum_{i\neq j} \rho_{ij} = \widetilde{\rho},$$

where $\widetilde{\rho} \in (0;1)$. *Then, if the additional condition*

$$\lim_{N\to\infty} \frac{1}{N(N-1)} \sum \sum_{i\neq j} (\rho_{ij} - \widetilde{\rho})^2 = 0$$

holds, the asymptotic distribution of the combined statistic $t(\widehat{\rho}^*, \kappa)$ *from (3.4) is not affected as long as* $\lambda_i = \lambda$.

The proof is given in the Appendix. Proposition 3.1 can be used, for instance, to justify the use of the modified inverse normal method when a finite number of "outliers" is present in the covariance matrix. Also, correlation matrices gradually approaching the constant correlation specification as $N \to \infty$ are allowed for. Although Proposition 3.1 assumes constant weights, the result can be extended to cover cases of variable weights, $\lambda_i \neq \lambda_j$, but under additional conditions, such as $\lambda_i \to \lambda$ as $N \to \infty$. There is controversy, however, as to whether one should allow for different weights in multi-unit studies. Thus, not allowing for variable weights across the units is not critical in practical applications.

The condition $\widetilde{\rho} > 0$ is crucial for the result. Should $\widetilde{\rho}$ tend to zero as N grows, one would conclude that units are uncorrelated, although cross-correlation may exist, even if not as strong as implied by the assumption $\rho_{ij} = \rho$.

For the case $\widetilde{\rho} = 0$, other parameterizations of the covariance matrix could be imposed. For instance, one could assume a geometrically decreasing specification,

$$Cov(t_i, t_j) = \rho^{|i-j|},$$

on the grounds that neighbouring units should exhibit the same correlation. However, one has to have a criterion by which units are sorted when specifying the covariance matrix this way. Obviously, the variance of the combined test statistic would have to be modified accordingly, just like the inference for the parameter ρ. Specifically, one could exploit the similarity to the autocovariance matrix of an autoregressive process of order 1, and, given an ordering, regress t_i on t_{i-1}, $i = 2, \ldots, N$. One should, however, not resort to different parameterization of the covariance matrix of the probits unless there are convincing reasons to do so.

As has been stressed before, the probits follow $N(0,1)$ distributions individually. Unfortunately, this does not imply multivariate normality of the probits, which would in turn imply asymptotic normality of $t(\widehat{\rho}^*, \kappa)$ under H_0. Finding a random vector with marginal normal distributions but whose distribution is not multivariate normal is a textbook problem. Let for instance (example 18.10 from Heike and Târcolea, 2000) $X' = (X_1; X_2)$ be a random vector with joint probability density function (pdf)

$$f(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}\left(x_1^2 + x_2^2\right)} \cdot \left[1 + x_1 \cdot x_2 \cdot e^{-\frac{1}{2}\left(x_1^2 + x_2^2 - 2\right)}\right], (x_1, x_2) \in \mathbb{R}^2.$$

As it can be easily checked, the marginal pdf of X_1 is

$$f(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2},$$

hence the marginal normality assumption holds. For reasons of symmetry, the same is true for X_2. The example illustrates the fact that, although the marginals are normal, the joint density is not[4]. Moreover, the normality of the sum of (weighted) normal distributions is only guaranteed for elements from a multivariate normal distribution. Continuing the example, the sum $U = X_1 + X_2$ is not normal, because its pdf is

$$\begin{aligned} f(u) &= \int_{-\infty}^{\infty} f(x, u - x)\, dx = \\ &= \frac{1}{\sqrt{4\pi}} e^{-\frac{u^2}{4}} + \frac{1}{\sqrt{128\pi}} \left(u^2 - 1\right) e^{-\frac{u^2}{2}}. \end{aligned}$$

Under what conditions does the inverse normal method map the original test statistics $S_1, \ldots, S_N$ to a multivariate normal distribution of the probits? A necessary and sufficient condition can be achieved by using the copula approach, which allows one to deal with the margins and the dependence structure separately - see Nelsen (1998) for an introduction to copulas. A copula

[4] For another example see Bickel and Doksum (2001, p. 533).

is a function that joins univariate distributions functions to form multivariate distribution functions. In fact, a copula can be viewed as a multivariate distribution function with univariate uniform margins restricted to the N-dimensional unit cube. Sklar's Decomposition Theorem (given for instance in Nelsen, 1998, p. 18) states that for a given N-dimensional distribution function and the corresponding N marginal distributions, there is a unique copula[5] linking them. This implies the known corollary that any monotonically increasing univariate transformation of the margins leaves the copula unchanged. Hence, the following Proposition follows immediately.

Proposition 3.2 *The test statistic $t(\rho)$ from (3.3) is $N(0,1)$ for any $\lambda_i \in \mathbb{R}$, $\sum_{i=1}^{N} \lambda_i \neq 0$ if and only if the statistics S_i that define t_i in equation (3.1) have the copula of a multivariate normal distribution.*

A sufficient (but not necessary) condition for the probits $t_1, \ldots, t_N$ to follow a multivariate normal distribution is obviously the multivariate normality of $S_1, \ldots, S_N$.

In some practical situations, a Central Limit Theorem leading to a multivariate normal distribution of the test statistics can be invoked. Stationarity of the stochastic processes used to model observed time-series typically leads to (multivariate) normal statistics. One should be aware, though, that stationarity *per se* is neither necessary, nor sufficient: there are cases of test statistics with non-standard distributions in stationary panels (such as the stationarity test due to Kwiatkowski et al., 1992), as well as examples of normal distributions arising from models with integrated processes (like the fractional LM integration test of Robinson, 1994, and Tanaka, 1999).

But even if the condition in Proposition 3.2 is violated, one still may be better off by correcting for cross-dependence than by ignoring it, since (strong) cross-correlation arguably causes more damage to a combined test statistic than deviations from normality of single test statistics. By means of Monte Carlo simulations, this claim is quantified in the following section for the case of unit root testing.

3.3 Panel tests

We first consider an example where a Central Limit Theorem assures a normal copula and hence satisfies the condition in Proposition 3.2. Then, we deal with

[5]This holds with the supplementary assumption that the margins are continuous, which is true in the case of the normal.

the already mentioned case of unit root testing with asymptotic non-normal distribution (Dickey-Fuller). The simulations were performed by means of GAUSS with 25,000 replications.

3.3.1 Standard case

The true data generating process (DGP) is

$$\begin{aligned} y_{it} &= \mu_i + \varepsilon_{it},\ i = 1, ..., N;\ t = 1, ..., T, \text{ where} \\ \varepsilon_t &= (\varepsilon_{it})_{i=1,\dots,N} \sim iiN(\mathbf{0}; \Sigma), \end{aligned}$$

with covariance matrix Σ first given by (3.2) as

$$\Sigma_1 = \begin{pmatrix} 1 & \rho_Y & \rho_Y & \cdots & \rho_Y & \rho_Y \\ \rho_Y & 1 & \rho_Y & \cdots & \rho_Y & \rho_Y \\ \rho_Y & \rho_Y & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \rho_Y & \rho_Y \\ \rho_Y & \rho_Y & \cdots & \rho_Y & 1 & \rho_Y \\ \rho_Y & \rho_Y & \cdots & \rho_Y & \rho_Y & 1 \end{pmatrix}.$$

For each unit, one computes the t statistic for the null hypothesis that the longitudinal mean is zero:

$$H_0 : \mu_i = 0,\ i = 1, \dots, N,$$

against the alternative

$$H_1 : \exists i \text{ } with \text{ } \mu_i < 0.$$

The significance levels are combined by means of the modified inverse normal method. Throughout this section, the weights in (3.4) are chosen as $\lambda_i = 1$ and[6] $\kappa = \kappa_1 = 0.2$. The p-values are calculated using the numerical cdf of the t distribution with $T - 1$ degrees of freedom, supplied by GAUSS. Probits are calculated using the numerical inversion of the standard normal cdf also available in GAUSS. Due to symmetry reasons, the covariance matrix of the probits t_i exhibits constant cross-correlation.

For Σ_1 with $\rho_Y = 0.2$, $\rho_Y = 0.5$ and $\rho_Y = 0.8$, the rejection frequencies under the null hypothesis are given in Tables 3.1 through 3.3.

For relatively small cross-correlation ($\rho_Y = 0.2$), the panel test is oversized

[6] All simulations were also carried out with $\kappa = \kappa_2$. The results were similar to those run with κ_1 and are therefore not reported here.

Table 3.1: Panel t test for Σ_1, $\rho_Y = 0.2$

	$T=50$			$T=100$			$T=250$		
N	1%	5%	10%	1%	5%	10%	1%	5%	10%
2	2.08	5.47	9.63	2.02	5.44	9.73	2.08	5.46	9.62
5	2.41	5.78	9.57	2.43	5.75	9.58	2.36	5.71	9.62
10	2.60	6.29	10.45	2.53	5.98	9.90	2.62	6.06	9.94
25	2.97	6.94	11.12	2.90	6.66	10.72	2.84	6.50	10.71

Note: Experimental size of $t(\widehat{\rho}^*, \kappa)$ from (3.4) with $\lambda_i = 1$ at nominal 1%, 5% and 10% level. Further information is given in the text.

Table 3.2: Panel t test for Σ_1, $\rho_Y = 0.5$

	$T=50$			$T=100$			$T=250$		
N	1%	5%	10%	1%	5%	10%	1%	5%	10%
2	1.82	5.50	10.13	1.78	5.46	10.19	1.58	5.25	9.75
5	1.91	5.79	10.21	1.93	5.64	10.06	1.93	5.75	10.38
10	1.80	5.65	10.43	1.74	5.57	10.28	1.69	5.64	10.55
25	1.22	5.21	10.22	1.43	5.27	10.13	1.35	5.41	10.16

Note: See Table 3.1.

at the 1% level. A better fit could be obtained for a different value of κ. The size distortions at the 5% level are of lower magnitude than those at the 1% level, and somewhat increase with growing N. The sizes improve with growing ρ_Y; the time dimension has limited relevance to the quality of the panel t test.

Continuing this set of experiments, we chose the covariance matrix Σ_2 as follows:

$$\Sigma_2 = \begin{pmatrix} 1 & \rho_Y & \rho_Y^2 & \cdots & \rho_Y^{N-2} & \rho_Y^{N-1} \\ \rho_Y & 1 & \rho_Y & \cdots & \rho_Y^{N-3} & \rho_Y^{N-2} \\ \rho_Y^2 & \rho_Y & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \rho_Y & \rho_Y^2 \\ \rho_Y^{N-2} & \rho_Y^{N-3} & \cdots & \rho_Y & 1 & \rho_Y \\ \rho_Y^{N-1} & \rho_Y^{N-2} & \cdots & \rho_Y^2 & \rho_Y & 1 \end{pmatrix}.$$

This specification leads to violation of the assumptions of Proposition 1. For $\rho_Y = 0.5$, the results are given in Table 3.4.

In this case, distortions occur, especially at the 1% (oversized) and 10% (undersized) level. They worsen with growing N. For up to $N = 10$ units, the size properties at the 5% and 10% levels are still acceptable.

Table 3.3: Panel t test for Σ_1, $\rho_Y = 0.8$

	$T=50$			$T=100$			$T=250$		
N	1%	5%	10%	1%	5%	10%	1%	5%	10%
2	1.16	5.15	10.22	1.15	4.92	9.88	1.05	5.19	10.03
5	1.04	5.12	10.05	1.09	4.88	9.82	1.00	4.86	9.99
10	0.96	4.68	9.60	0.99	4.83	9.80	1.00	4.84	9.50
25	1.04	4.92	9.73	0.93	4.79	9.62	0.99	4.63	9.72

Note: See Table 3.1.

Table 3.4: Panel t test for Σ_2, $\rho_Y = 0.5$

	$T=50$			$T=100$			$T=250$		
N	1%	5%	10%	1%	5%	10%	1%	5%	10%
2	1.64	5.31	10.26	1.70	5.42	10.13	1.70	5.75	10.21
5	2.58	6.18	10.13	2.55	6.08	10.10	2.44	5.83	9.95
10	2.35	5.52	9.22	2.55	6.05	9.78	2.64	5.97	9.42
25	1.87	4.92	8.21	1.67	4.66	7.72	1.75	4.71	7.93

Note: See Table 3.1.

3.3.2 Non-standard case

The true DGP changes to

$$\begin{aligned} y_{it} &= \sum_{j=1}^{t} \varepsilon_{ij}, \; i = 1, ..., N; \; t = 1, ..., T, \text{ where} \\ \varepsilon_t &= (\varepsilon_{it})_{i=1,...,N} \sim iiN(\mathbf{0}; \Sigma). \end{aligned}$$

We run for N units with T observations each a Dickey-Fuller (DF) test for a unit root. The p-value for each unit is calculated based on simulated cdfs with 100,000 replications for the given sample size[7]. The estimated DF regression is

$$\Delta y_{it} = \widehat{a}_i y_{it-1} + \widehat{\varepsilon}_{it},$$

where the null hypothesis

$$H_0 : a_i = 0, \;\; i = 1, \ldots, N$$

is tested against the alternative

$$H_1 : \exists i \; with \; a_i < 0.$$

[7]We waived the possibility of using the methods proposed by Adda and Gonzalo (1996) and MacKinnon (1996), since they also work with simulated p-values, but ingeniously interpolated. Therefore, no accuracy is lost using an empirical cdf of the Dickey-Fuller test statistic to calculate p-values.

In addition to the DF test, we performed the nonlinear instrument variables (NIV) unit root test by Chang (2002). The test regression is

$$y_{it} = \widetilde{b}_i y_{it-1} + \widetilde{\varepsilon}_{it},$$

but estimation is done with the instrument z_{it-1} for y_{it-1}, where $z_{it-1} = F(y_{it-1})$, with $F(u) = ue^{-c_i|u|}$. The constants c_i are defined by Chang as $c_i = 3T^{-1/2}s^{-1}(\Delta y_{it})$, where $s^2(\Delta y_{it}) = T^{-1}\sum_{t=2}^{T}(\Delta y_{it})^2$. The null hypothesis changes to

$$H_0 : b_i = 1, \;\; i = 1, \dots, N,$$

and is tested against the alternative

$$H_1 : \exists i \; with \; b_i < 1.$$

The combination of p-values due to Fisher[8], advocated by Maddala and Wu (1999), is included as benchmark. We do not include the PANIC approach due to Bai and Ng (2004) in this simulation study, due to the fact that the DGP they work with is different from ours: they assume the existence of at least a common factor, and the covariance matrix of the innovations is only allowed to exhibit weak correlation.

Table 3.5: Panel unit root test for Σ_1, $\rho_Y = 0.2$

		$T=50$			$T=100$			$T=250$		
	N	1%	5%	10%	1%	5%	10%	1%	5%	10%
Fisher	2	1.18	5.43	10.53	1.14	5.36	10.22	1.04	5.23	10.50
	5	1.19	5.40	10.45	1.16	5.73	11.04	1.14	5.35	10.26
	10	1.36	5.81	11.04	1.29	5.88	11.26	1.17	5.80	11.32
	25	1.89	7.14	13.17	1.74	7.22	13.06	1.66	7.21	13.34
$t(\widehat{\rho}^*, \kappa)$	2	2.00	5.22	9.12	1.79	4.90	8.64	2.04	5.41	9.24
	5	1.28	3.93	6.92	1.21	3.88	6.98	1.32	4.17	7.36
	10	0.80	2.92	5.50	0.74	3.01	5.85	0.79	3.18	5.99
	25	0.29	1.84	4.07	0.27	1.91	4.21	0.34	2.10	4.68
NIV	2	1.69	5.97	10.94	1.32	5.51	10.31	1.37	5.69	10.29
	5	1.37	5.28	9.63	1.20	5.09	9.38	1.17	4.90	9.42
	10	1.22	5.02	9.28	1.27	4.89	9.12	1.21	4.78	9.10
	25	1.39	5.02	9.03	1.30	4.50	8.28	1.13	4.55	8.07

Note: Experimental size of Maddala/Wu panel unit root test (Fisher), of $t(\widehat{\rho}^*, \kappa)$ from (3.4), and of the nonlinear instrumental test (NIV) due to Chang at nominal 1%, 5% and 10% level. Units are equally weighted. Further information is given in the text.

The results of combining the corresponding probits, together with the re-

[8]Fisher suggests the use of logarithmic transformations of p-values, instead of applying the inverse normal method. This produces an overall test statistic following a $\chi^2(2N)$ distribution, and the null hypothesis is rejected for large positive values. Although it unrealistically assumes independent units, the behaviour of this panel unit root test is of interest in econometric practice, being implemented in EViews from version 5 onwards.

Table 3.6: Panel unit root test for Σ_1, $\rho_Y = 0.5$

		$T=50$			$T=100$			$T=250$		
	N	1%	5%	10%	1%	5%	10%	1%	5%	10%
Fisher	2	1.12	5.36	10.67	1.21	5.50	10.78	1.15	5.49	10.69
	5	1.80	6.98	12.80	1.76	7.30	13.02	1.72	6.98	12.52
	10	3.06	9.96	16.18	3.03	9.89	16.18	3.16	10.07	16.11
	25	6.90	16.62	23.68	7.44	17.08	24.15	7.28	16.76	23.91
$t(\widehat{\rho}^*, \kappa)$	2	1.88	5.14	9.30	2.08	5.62	9.72	1.75	4.98	9.25
	5	2.02	5.23	9.01	2.06	5.28	8.78	1.77	4.95	8.48
	10	1.93	5.32	8.86	1.76	5.19	8.74	1.73	4.79	8.13
	25	1.87	5.20	8.64	2.00	5.34	8.81	1.63	4.64	8.04
NIV	2	1.82	6.40	11.28	1.89	6.53	11.73	1.52	6.38	11.42
	5	2.58	8.06	13.65	2.34	7.67	13.12	2.09	7.90	13.30
	10	3.78	10.70	16.57	3.57	10.43	16.24	3.27	9.77	15.66
	25	8.04	16.79	22.92	7.57	16.44	22.80	7.18	15.76	22.16

Note: See Table 3.5.

Table 3.7: Panel unit root test for Σ_1, $\rho_Y = 0.8$

		$T=50$			$T=100$			$T=250$		
	N	1%	5%	10%	1%	5%	10%	1%	5%	10%
Fisher	2	1.76	6.80	12.37	1.72	6.82	12.24	1.82	6.82	12.12
	5	4.64	12.25	18.54	4.41	12.15	18.42	4.82	12.56	18.78
	10	9.66	18.84	24.64	9.82	18.78	24.54	10.25	19.20	25.03
	25	20.85	28.74	33.24	20.92	28.42	33.04	20.52	28.12	32.63
$t(\widehat{\rho}^*, \kappa)$	2	1.89	5.55	10.01	2.05	5.72	10.35	2.04	5.68	10.37
	5	2.72	6.06	10.50	2.63	6.05	10.37	2.75	6.40	10.73
	10	3.05	6.68	10.96	2.95	6.58	11.10	2.82	6.42	10.96
	25	3.26	6.79	11.26	3.36	6.81	11.34	3.10	6.44	11.08
NIV	2	2.78	8.60	14.52	2.68	8.70	14.47	2.59	8.86	14.69
	5	7.32	16.07	22.37	6.46	15.48	21.90	6.50	15.54	22.13
	10	14.37	23.98	29.82	14.06	24.15	30.24	13.97	23.80	29.90
	25	26.00	33.69	37.89	26.47	33.86	37.94	26.14	33.88	38.02

Note: See Table 3.5.

sults of the NIV test, are given for simulations under the null hypothesis in Tables 3.5, 3.6, and 3.7, respectively, for $\Sigma = \Sigma_1$ with ρ_Y equal to 0.2, 0.5 and 0.8.

We learn that the experimental level very much depends on the cross-correlation ρ_Y. The Fisher test is increasingly oversized with growing ρ_Y and N, while the size distortions do not depend on the sample size T. Compared to that, the modified inverse normal method has very good size properties at the 5% level for all N and T, especially for medium and strong correlation. For weak correlation, $\rho_Y = 0.2$, the modified inverse normal method delivers an undersized test and is not to be used with more than 5 units. The worst distortions occur, as in the previous subsection, for the 1% level, irrespective of the magnitude of the cross-correlation. The NIV test performs roughly like the Fisher test; at the 5% level, its size properties are good only in case of weak cross-correlation (the same observation was made by Im and Pesaran, 2003).

Further, we wish to examine how the tests behave if the constant correlation assumption is violated as well, but the deviations are likely to be tolerable under the conditions of Proposition 3.1. We report in the following the mean size of the tests over 20 randomly generated covariance matrices[9]. To reduce computational burden, only 5,000 Monte Carlo replications were conducted for each studied case. The covariance matrices are drawn following the recipe used by Chang (2002) for her simulation studies. Specifically, she uses the spectral decomposition, where the eigenvectors are generated as $H = M(M'M)^{-1/2}$, with M a $N \times N$ matrix with elements from a uniform distribution on the interval $[0; 1]$, and the eigenvalues are randomly drawn from a uniform distribution on $[0.1; 1]$. We, however, implement a slight modification. In order to avoid criticism of the type expressed by Im and Pesaran (2003), who argue that the covariance matrices generated by this recipe have low mean correlation, we use the eigenvalues of Σ_1 with $\rho_Y = 0.5$, of which the largest is $(N+1)/2$. The results are given in Table 3.8.

Table 3.8: Panel unit root test for random correlation matrices

		$T = 50$			$T = 100$			$T = 250$		
	N	1%	5%	10%	1%	5%	10%	1%	5%	10%
Fisher	2	1.10	5.30	10.48	1.12	5.25	10.23	1.07	5.17	10.35
	5	1.47	6.12	11.59	1.40	6.11	11.47	1.40	6.15	11.32
	10	2.28	8.14	14.02	2.12	7.79	13.57	2.15	7.95	13.90
	25	4.91	12.76	19.35	4.69	12.50	18.94	4.93	12.82	19.65
$t(\hat{\rho}^*, \kappa)$	2	1.96	5.12	9.04	1.96	5.06	8.98	1.99	5.21	9.10
	5	1.71	4.67	7.99	1.63	4.49	7.92	1.68	4.61	7.95
	10	1.29	4.15	7.36	1.31	4.11	7.30	1.29	4.22	7.54
	25	1.24	4.28	7.44	1.21	4.32	7.67	1.32	4.37	7.67
NIV	2	1.61	6.17	10.96	1.47	5.82	10.60	1.43	5.86	10.57
	5	1.99	6.92	11.84	1.84	6.52	11.34	1.74	6.46	11.34
	10	2.59	8.04	13.22	2.32	7.57	12.62	2.11	7.17	12.30
	25	4.42	11.14	16.75	4.28	11.11	16.63	3.97	10.53	16.03

Note: See Table 3.5.

We observe the Fisher test to be (again) oversized; the NIV test, while somewhat improving on the size problem, does not solve it. Opposed to that, the modified inverse normal method behaves particularly well at the 5% level; the test is somewhat oversized at the 1% level, and conservative at the 10% level.

[9]The maximal and minimal size over the 20 matrices were not essentially different – the differences to the mean size are no larger than 25% – and are available upon request from the authors.

3.4 Concluding remarks

The idea of combining p-values can be of much use in panel analyses. E.g., Maddala and Wu (1999) and Choi (2001) proposed the combination of significance levels when testing for unit roots in panel data. In particular, Choi (2001) favored the so-called inverse normal method. For this method, Hartung (1999) suggests a modification that allows for constant correlation between test statistics.

We first extend the scope of the modified inverse normal method by proving robustness to more general correlation matrices. Second, a necessary and sufficient condition for its validity is provided, namely: a normal copula of the original test statistics. A sufficient condition for that clearly is multivariate normality of the statistics.

As an application, we study experimentally the panel t test with normal iid time observations, which turns out to work especially well in the case of large cross-correlation. The property of a normal copula does not hold with Dickey-Fuller unit root tests. We provide simulation results that quantify the behaviour of the modified inverse normal method when applied to Dickey-Fuller tests from cross-correlated panels. It is found that, for medium and strong cross-correlation, the advocated combination of p-values delivers very good results at the 5% level. Otherwise, although less desirable, the results are in most situations superior to ignoring the cross-correlation problem.

Appendix

Proof of Proposition 3.1

The combination of the probits is

$$t\left(\widehat{\rho}^{*},\kappa\right)=\frac{\sum_{i=1}^{N}\lambda t_i}{\sqrt{N\lambda^2\left(1-\widehat{\rho}\right)+\lambda^2N^2\widehat{\rho}}}\,.$$

For simplicity, assume that $\widehat{\rho}$ takes admissible values a.s., and ignore the small-sample correction. For the numerator, it holds

$$Var\left(\sum_{i=1}^{N}\lambda t_i\right)=\lambda^2\tau'\Sigma\tau,$$

where $\tau' = (1, 1, \ldots, 1)_{1\times N}$, or

$$Var\left(\sum_{i=1}^{N} \lambda t_i\right) = \lambda^2 \sum_{i=1}^{N}\sum_{j=1}^{N} \rho_{ij}.$$

Since, as assumed, $Var\left(\sum_{i=1}^{N} \lambda t_i\right) = O\left(N^2\right)$, we write

$$t\left(\widehat{\rho}\right) = \frac{\frac{1}{N}\sum_{i=1}^{N} \lambda t_i}{\sqrt{\frac{1}{N^2}\left(N\lambda^2\left(1-\widehat{\rho}\right) + \lambda^2 N^2 \widehat{\rho}\right)}},$$

to obtain well defined limiting quantities for numerator and denominator. It then holds

$$\lim_{N\to\infty} Var\left(\frac{1}{N}\sum_{i=1}^{N} \lambda t_i\right) = \lambda^2 \widetilde{\rho}.$$

Let us now examine the denominator of the combined statistic. According to Hartung's method, one first obtains an estimator for ρ,

$$\widehat{\rho} = 1 - \frac{1}{N-1}\sum_{i=1}^{N}\left(t_i - \overline{t}\right)^2 = 1 - \frac{\mathbf{t}' K \mathbf{t}}{N-1},$$

where $K = I - \frac{1}{N}\tau\tau'$ is idempotent and $tr(K) = N - 1$. By using well-known properties of quadratic forms, the expected value of $\widehat{\rho}$ is shown to be

$$\begin{aligned}
E\left(\widehat{\rho}\right) &= 1 - tr\left(\frac{K\Sigma}{N-1}\right) = 1 - \frac{1}{N-1} tr\left(\Sigma - \frac{1}{N}\tau\tau'\Sigma\right) \\
&= 1 - \frac{N}{N-1} + \frac{1}{N\left(N-1\right)} tr\left(\tau\tau'\Sigma\right) = -\frac{1}{N-1} + \frac{1}{N\left(N-1\right)} tr\left(\tau'\Sigma\tau\right) \\
&= -\frac{1}{N-1} + \frac{1}{N\left(N-1\right)}\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{ij} = \frac{1}{N\left(N-1\right)}\sum\sum\nolimits_{i\neq j}\rho_{ij}.
\end{aligned}$$

It hence follows

$$\lim_{N\to\infty} E\left(\frac{1}{N^2}\left(N\lambda^2\left(1-\widehat{\rho}\right) + \lambda^2 N^2 \widehat{\rho}\right)\right) = \lambda^2 \widetilde{\rho}.$$

The variance of the estimator $\widehat{\rho}$ is given by

$$Var\left(\widehat{\rho}\right) = 2tr\left(\frac{K\Sigma K\Sigma}{\left(N-1\right)^2}\right),$$

which is now shown to disappear for $N \to \infty$. To this purpose, one can decompose

$$\Sigma = \Sigma_{\widetilde{\rho}} + \Sigma_N,$$

where $\Sigma_{\widetilde{\rho}}$ is a correlation matrix with constant off-diagonal elements $\widetilde{\rho}$. By calculating the multiplication

$$K\Sigma K\Sigma = \left(I - \frac{1}{N}\tau\tau'\right)\left(\Sigma_{\widetilde{\rho}} + \Sigma_N\right)\left(I - \frac{1}{N}\tau\tau'\right)\left(\Sigma_{\widetilde{\rho}} + \Sigma_N\right),$$

and using, as Hartung, the relationship

$$K\Sigma_{\widetilde{\rho}} = (1 - \widetilde{\rho})\,K,$$

the condition $Var\,(\widehat{\rho}) \to 0$ reduces to $tr\,(\Sigma_N^2) = o\,(N^2)$, which is equivalent to the second assumption of the proposition. By applying Slutsky's theorem, the denominator is shown to converge to the square root of the numerator variance. Hence, under the conditions of Proposition 3.2, the result follows.

Chapter 4

Combining Multi-Country Evidence on Unit Roots: The Case of Long-Term Interest Rates

4.1 Introduction

Panel unit root tests have recently attracted a lot of attention, see Banerjee (1999) or Baltagi and Kao (2000) for overviews. They were originally designed for multi-country studies and are intended to combine evidence from individual cross-sections to one joint significance level. We apply those tests to monthly interest rates and differentials to throw light on the interest rate linkage for OECD countries. For earlier evidence see e.g. Kirchgässner and Wolters (1995). Specifically, we use two approaches. First, we employ panel Dickey-Fuller (DF) tests, and second, we use methods that combine p values and can in principle be applied to any type of test. All procedures but one have to assume independent units, which is not very realistic for OECD countries. The exception is the so-called modified inverse normal method by Hartung (1999), which can handle cross-correlation. Recent tests accounting for cross-correlation are by Chang (2002), Bai and Ng (2004), Moon and Perron (2004), and Breitung and Das (2005). They are technically more complicated and due to space limitations not considered here.

The chapter is structured as follows. After the introduction, we briefly review the most common panel unit root tests. Section 4.3 surveys approaches that combine p values, and in particular the modified inverse normal method.

It also compares the methods by means of a small simulation exercise.The fourth section contains the empirical results, and Section 4.5 concludes.

4.2 Combining Dickey-Fuller tests

In this section we spell out the assumptions behind the most widely used panel DF tests, which are implemented for instance in *EViews.* The general setup assumes N countries (or units in general), $i = 1, \ldots, N$, and T_i time series observations in each country, $t = 1, \ldots, T_i$. All tests

- rely on sequential limit theory, first $T_i \to \infty$, then $N \to \infty$,
- assume independent units,
- and, as a consequence, result in limiting normal distributions.

The general time series model is

$$x_{i,t} = d_{i,t} + \alpha_i \, x_{i,t-1} + e_{i,t} \, ,$$

where $d_{i,t}$ denotes a deterministic component and $e_{i,t}$ a stationary and invertible ARMA process. Throughout, those components are individual-specific.

The first test by Levin, Lin and Chu (2002) (or LLC for short) imposes a common root restriction ($\alpha_1 = \cdots = \alpha_N = \alpha$) under the null as well as under the alternative hypothesis:

$$H_0 : \ \alpha = 1 \, , \qquad H_1 : \ \ |\alpha| < 1 \, .$$

The idea of their approach is to adjust all series individually for deterministics and short-run autocorrelation, and then to perform a pooled DF test with the residuals. The pooled DF statistic has to be adjusted for a bias term in order to converge to a limiting standard normal distribution. Breitung (2000) argues that this bias correction may result in a substantial loss of power. He proposes a modification that circumvents bias adjustment. It amounts to subtracting the first observation instead of demeaning the series. This test will be called LLCB in the following.

Im, Pesaran and Shin (2003) (or IPS for short) criticize the assumption of a common root under the alternative. Instead they allow for heterogeneity of all parameters. To maintain power they require $|\alpha_i| < 1$ for a sufficiently large number of units. Consequently, all parameters are individual-specific, and it is natural to perform N tests individually, and to average over individual DF statistics.

4.3 Combining p values

The idea of combining p values of several statistics to obtain a joint level of significance can be traced back to R.A. Fisher, see, for instance, Fisher (1954, p. 100). Let S_i $(i = 1, ..., N)$ denote continuous test statistics for the null hypotheses

$$H_i^{(0)} : \ \theta_i = \theta_i^{(0)} , \quad i = 1, \ldots, N ,$$

against the one-sided alternatives

$$H_i^{(1)} : \theta_i < \theta_i^{(0)} .$$

Let p_i denote the corresponding p values. The so-called probits t_i are defined as quantiles,

$$t_i = \Phi^{-1}(p_i)$$

with Φ^{-1} being the inverse of the standard normal distribution function. By construction, the p values follow a uniform law on $[0, 1]$, hence

$$-2 \ln(p_i) \ \sim \ \chi^2(2) \quad \text{and} \quad t_i \ \sim \ \mathcal{N}(0, 1) .$$

The original proposal by Fisher combines the χ^2 distributions,

$$F = -2 \sum_{i=1}^{N} \ln(p_i) , \tag{4.1}$$

while the so-called inverse normal method relies on summing normalized probits:

$$t = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N}} . \tag{4.2}$$

Maddala and Wu (1999) suggest the use of F when testing for unit roots from a multi-country panel, while t is favoured by Choi (2001). Under the joint null hypothesis

$$H^{(0)} : \ \theta_i = \theta_i^{(0)} \quad \text{for all } i = 1, \ldots, N$$

and under the assumption of independent units the following holds in finite samples:

$$F \ \sim \ \chi^2(2N) \quad \text{and} \quad t \ \sim \ \mathcal{N}(0, 1) .$$

The tests are one-sided. Fisher's test rejects for too large values, while t is significant for too small values. The advantages of combining p values are the applicability

- with all type of test problems where p values are available,
- in the presence of unbalanced panels,
- with standard distributions without complicated tabulations.

The major drawback of course is that independence of units still has to be assumed.

We now briefly present the modification of the inverse normal method proposed by Hartung (1999) to account for correlation between the units. In fact, he assumes constant correlation between the probits,

$$corr(t_i, t_j) = \rho \quad \text{for } i \neq j, \quad i, j = 1, ..., N. \tag{4.3}$$

If ρ were known, we could generalize (4.2) in obvious manner:

$$t(\rho) = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N + N\,(N-1)\,\rho}}, \tag{4.4}$$

with $Var\left(\sum_{i=1}^{N} t_i\right) = N + N\,(N-1)\,\rho$. For empirical applications ρ has to be estimated,

$$\widehat{\rho} = 1 - Q \quad \text{with } Q = \frac{1}{N-1} \sum_{i=1}^{N} \left(t_i - \bar{t}\right)^2$$

where $\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_i$. Hartung (1999) proves that the quadratic form Q relates to a χ^2 distribution as follows:

$$\frac{N-1}{1-\rho} Q \sim \chi^2(N-1).$$

Consequently, $\widehat{\rho}$ is an unbiased and consistent estimator:

$$E(\widehat{\rho}) = \rho, \quad Var(\widehat{\rho}) = \frac{2\,(1-\rho)^2}{N-1}.$$

For applied work in finite samples the estimator is modified yet again:

$$\widehat{\rho}^* = \max\left(-\frac{1}{N-1}, \widehat{\rho}\right).$$

The lower bound $-\frac{1}{N-1}$ avoids negative variance terms when replacing ρ in the denominator of (4.4) by $\widehat{\rho}^*$. To guarantee strictly positive variances Hartung (1999) introduces the parameter $\kappa > 0$ and suggests

$$t(\widehat{\rho}^*) = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N + N\,(N-1)\,\left[\widehat{\rho}^* + \kappa\sqrt{\frac{2}{N+1}}\,(1-\widehat{\rho}^*)\right]}}\,. \tag{4.5}$$

On experimental grounds Hartung (1999) proposes $\kappa = 0.2$. Given consistency (as $N \to \infty$) of the estimator $\widehat{\rho}^*$, Hartung's modified inverse normal method test statistic is again compared with the approximating standard normal distribution.

Finite sample p values required for unit root testing can be computed from MacKinnon (1996) and are available e.g. in *EViews*. Hence, a further item can be added to the list of the advantages of the modified inverse normal method: simplicity.

The approach by Hartung (1999) has been further investigated by Demetrescu, Hassler and Tarcolea (see Chapter 3). First, they study a violation of (4.3) and prove robustness against correlation varying at a certain degree. Second, they emphasize that marginal normality of the probits is not sufficient for limiting normality of $t(\widehat{\rho}^*)$. Third, they provide simulation evidence that the modified inverse normal method is reasonably reliable when applied to ADF tests in correlated panels. Here, we add some experimental evidence relevant to the current situation: Under the null hypothesis of N integrated series, they are assumed to be cointegrated of rank $N-1$. Under the alternative hypothesis all series are I(0). We try to mimic an empirical situation that is plausible to be encountered with monthly international interest rates:

$$\Delta x_{1,t} = \phi\, x_{1,t-1} + \varepsilon_{1,t}, \quad t = 1, \ldots, T = 200, \tag{4.6}$$

$$\Delta x_{i,t} = -0.5\,(x_{i,t-1} - x_{1,t-1}) + \varepsilon_{i,t}, \quad i = 2, \ldots, N = 10\,. \tag{4.7}$$

The innovations are assumed to be serially independent Gaussian with constant contemporaneous correlation:

$$\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{N,t})' \;\sim\; ii\mathcal{N}(0,\, \Sigma),$$

$$\text{with } \Sigma = \begin{pmatrix} 1 & \rho_\varepsilon & \cdots & \rho_\varepsilon \\ \rho_\varepsilon & 1 & \cdots & \rho_\varepsilon \\ \vdots & \vdots & \ddots & \vdots \\ \rho_\varepsilon & \rho_\varepsilon & \cdots & 1 \end{pmatrix}.$$

Then ADF tests with intercept (without trend) were applied to $x_{i,t}$ with data-driven lag length selection, and p values were combined as in (4.5). For comparison we also include the test proposed by Maddala and Wu (1999) relying on (4.1). The simulations were performed by means of GAUSS with 5,000 replications. Table 4.1 contains the rejection frequencies when testing at the nominal sizes of 1%, 5% and 10%.

We observe that the rates of rejection under the null hypothesis of integration vary only little with ρ_ε. More interestingly, with only $N = 10$ cointegrated units the standard normal distribution provides a reasonable approximation for the finite sample distribution of Hartung's modified inverse normal test statistic, while Fisher's original proposal F without correction for cross-correlation is grossly oversized. From the right panel in Table 4.1 we further learn that the power of the $t(\widehat{\rho}^*)$ test decreases with growing ρ_ε.

Table 4.1: Size (and power) with cointegrated series

		$H_0 : \phi = 0$			$H_1 : \phi = -0.1$		
	$\rho_\varepsilon =$	0.2	0.5	0.8	0.2	0.5	0.8
$t(\widehat{\rho}^*)$	1%	2.44	1.94	1.26	67.86	49.26	34.76
	5%	7.82	7.18	5.94	91.84	83.08	71.98
	10%	14.50	12.74	11.06	97.56	93.58	87.06
F	1%	19.32	17.78	15.62	98.12	96.24	91.64
	5%	25.94	23.78	21.82	99.44	98.02	95.48
	10%	30.30	28.20	25.70	99.68	98.74	97.14

Note: Percentage of rejection for $N = 10$ units with $T = 200$ observations generated by (4.6) and (4.7). Further details are given in the text.

4.4 Empirical results

We use monthly data of nominal long-term interest rates[1] from 12 OECD countries: Austria, Canada, Finland, France, Germany, Italy, Japan, Netherlands, Portugal, Spain, United Kingdom, and Unites States. The sample starts in 1985.01 and ends in 2002.12, leaving us with $T = 216$ time series observations.

[1] Government bonds rates with 10 years to maturity.

All computations were performed with the help of *EViews*. Augmented Dickey-Fuller tests (ADF) were computed with automatic lag length selection relying on Schwarz' information criterion (SIC) and with a maximum lag length of 14. The estimation of long-run variance for the LLC test was carried out with the quadratic spectral kernel using Andrew's bandwidth selection. Throughout, $p_{(i)}$ denote ordered p values of individual ADF tests.

In order to assess whether long-term interest rates may be considered as integrated of order one, ADF tests with intercept but without linear trend are applied to the levels of the series. The outcome is summarized in Table 4.2. No individual test rejects at the 10% level, in fact the p values due to MacKinnon (1996) vary from 30.87% up to 80.70%. It is still interesting to study the outcome of combined results. Only the LLCB test is highly significant with a p value far below $p_{(1)}$, while, in contrast, the majority of tests produces levels above the maximum individual p value (0.807). The LLC test and the modified inverse normal method have values between the minimum and maximum p values. For the rest of the chapter we assume the considered interest rates to be integrated of order one. This is not rejected by Hartung's modification of the inverse normal method accounting for cross-correlation.

Table 4.2: p values testing interest rates

$p_{(1)}$	0.3087	*LLC*	0.4609	F	0.9895
$p_{(7)}$	0.6713	*LLCB*	0.0062	t	0.9216
$p_{(12)}$	0.8070	*IPS*	0.8875	$t(\hat{\rho}^*)$	0.6688

Note: Results are presented for individual ADF tests, the LLC , LLCB and IPS test. F, t, and $t(\hat{\rho}^*)$ are defined in (4.1), (4.2), and (4.5), respectively.

Now, we turn to the investigation of international interest rate differentials. Stationarity thereof is a necessary condition for the uncovered interest rate parity (UIP) hypothesis to hold in the presence of integrated exchange rates, see Wolters (2002) for a recent review of related theoretical and empirical literature. Since Kirchgässner and Wolters (1995) found that German and US interest rates influence other countries, cf. also Kirchgässner and Wolters (1993), we study differentials between all countries and the US and Germany. ADF tests are performed with intercept but without trend. We present p values for US differentials in the left panel of Table 4.3.

The differentials between the US and Germany, the Netherlands and Japan are significant at the 10% level. All other countries are partly very far from conventional significance levels. Joint panel tests ignoring cross-correlation transform this rather mixed evidence into significance around the 5% or 10%

Table 4.3: p values testing differentials

US				GERMANY			
$p_{(1)}$	0.0794	LLC	0.3538	$p_{(1)}$	0.0109	LLC	0.0775
$p_{(3)}$	0.0947	$LLCB$	0.0323	$p_{(3)}$	0.1248	$LLCB$	0.2089
$p_{(5)}$	0.1098	IPS	0.0297	$p_{(5)}$	0.2069	IPS	0.0365
$p_{(7)}$	0.2663	F	0.0606	$p_{(7)}$	0.4609	F	0.0616
$p_{(9)}$	0.6493	t	0.0265	$p_{(9)}$	0.5988	t	0.0352
$p_{(11)}$	0.7263	$t(\widehat{\rho}^*)$	0.1817	$p_{(11)}$	0.7321	$t(\widehat{\rho}^*)$	0.1716

Note: See Table 4.2.

level – with the exception of the LLC test, which is far from being significant. The modified inverse normal method produces a more conservative result not being indicative of significance.

Finally, consider the analogous results for differentials between the German and all other interest rates. Only the Netherlands and US are significant at the 5% and 10% levels, respectively. All other individual test statistics are again more or less far from conventional significance, see the right panel in Table 4.3. Joint panel tests melt this very heterogeneous evidence into significance at the 5% or 10% level – with the sole exception of the LLCB test and of the modified inverse normal method accounting for cross-correlation. Notice that the LLC test and Breitung's variant somehow contradict themselves. If we consider the differential between the US and Germany as stationary (individual p value of 7.94%) then the outcomes of the right and left panel in Table 4.3 should be concurring. This is in fact the case for all tests but LLC and LLCB.

4.5 Concluding remarks

We tested US and German interest rate differentials for nonstationarity using recent panel techniques. Unit root tests assuming a common autoregressive root over the countries seem to produce contradictory results with respect to stationarity of US and German differentials. The outcome of the IPS and Fisher tests and the inverse normal method are indicative in favour of the UIP hypothesis, no matter whether differentials with respect to US or German rates are tested. The modified inverse normal method, however, which is the only one accounting for cross-correlation between countries, produces p values of only 18% and 17% for US and German differentials, respectively. We conclude that ignoring or modelling cross-correlation in multi-country studies may heavily affect the outcome of nonstationary panel analyses.

Zusammenfassung

Paneldaten werden analysiert, um die Entwicklung ökonomischer (wie auch sozialer und technischer) Prozesse zu beschreiben, zu erklären, zu prognostizieren und zu kontrollieren. Da Paneldaten mehr Information als einzelne Zeitreihen beinhalten, werden sie benutzt, um Güte zu gewinnen, wenn man ökonomische Hypothesen testet, bzw. um Genauigkeit zu gewinnen, wenn man relevante Parameter schätzt. Allerdings verursachen Panels makroökonomischer Größen Schwierigkeiten, die durch die Natur der Paneldaten zustandekommen, und zwar durch die Kombination der Längsschnitt- und der Querschnittseigenschaften makroökonomischer Daten.

Einerseits gilt es zu beachten, dass die einzelnen Zeitreihen des jeweiligen Längsschnitts instationär sein können. Dies beeinflusst die Inferenz über die untersuchten Zusammenhänge zwischen den gegebenen makroökonomischen Größen. Von den vielen Formen von Instationarität sind zwei besonders wichtig: Zum einen Erwartungswertinstationarität in Form linearer Trends, und zum anderen Varianzinstationarität in Form so genannter integrierter Prozesse, die durch einen stochastischen Trend getrieben werden. Ein Prozess integriert der Ordnung Eins - I(1) - ist stationär nach einmaliger Differenzbildung. Äquivalent kann man sagen, dass ein solcher Prozess eine autoregressive Einheitswurzel aufweist.

Diese Eigenschaft der Instationarität (sei sie in Form deterministischer oder stochastischer Trends) ist insofern wichtig, da Scheinregressionen bei instationären Zeitreihen wesentlich öfter als bei stationären Zeitreihen auftauchen, eben wegen der Anwesenheit von Trends. Die Frage, ob ein Zusammenhang zwischen den untersuchten Größen besteht, geht in die Frage über, ob es einen oder mehrere gemeinsame Trends gibt. Im Falle integrierter Zeitreihen lässt sich diese Frage mit Hilfe der Kointegrationsanalyse beantworten. Dabei sind integrierte Zeitreihen genau dann kointegriert, wenn es eine Linearkombination gibt, die stationär ist. Alternativ lässt sich Kointegration als ein sogenanntes Fehlerkorrekturmodell darstellen, in dem die Inkremente zum Zeitpunkt

t derart von der Abweichung zum Zeitpunkt $t-1$ von dem langfristigen Gleichgewicht abhängen, dass das Gleichgewicht wieder hergestellt wird. Im Falle von Zeitreihen mit deterministischen Trends haben Hatanaka und Yamada (2003) den Begriff Ko-trending geprägt. Da viele makroökonomische Variablen, wie Zinssätze, Preisniveaus und Wechselkurse, als integriert von Ordnung Eins gelten, ist die empirische Relevanz dieser Eigenschaft offensichtlich.

Andererseits sind Abhängigkeiten zwischen den Paneleinheiten zu beachten, die ein beträchtlicher Störfaktor bei den Untersuchungen sein können: Z.B. weist die statistische Evidenz in einzelnen Paneleinheiten auf die Gültigkeit (bzw. Nichtgültigkeit) der untersuchten Hypothese hin, weil diese auf Panelebene gegeben ist, oder lediglich durch Zufall, der dadurch verstärkt wird, dass die Einheiten untereinander abhängig sind? Solche Abhängigkeiten sind insbesondere bei regionalen Datensätzen zu erwarten, bei denen die Beobachtungen von aneinander angrenzenden Regionen (Ländern) in der Regel hoch korreliert sind.

Dieses Problem der Kreuz-Abhängigkeit in Panels ist aber genau bei instationären Daten besonders schwierig zu behandeln, weil es schwerwiegende Verzerrungen zur Folge haben könnte. Beispielsweise wurde es in der ersten Generation von Paneleinheitswurzeltests völlig ignoriert, siehe Levin, Lin und Chu (2002) oder Im, Pesaran und Shin (2002). Deshalb stellt die Berücksichtigung von Korrelationen oder sogar Kointegration zwischen den Querschnittseinheiten einen wichtigen Schwerpunkt bei der Weiterentwicklung der Paneldatenmethoden dar. Tests, die eine kontemporäre Korrelation zulassen, werden der zweiten Generation von Paneleinheitswurzeltests zugeordnet (vgl. Breitung und Pesaran, 2006). Dabei werden zwei unterschiedliche Strategien verfolgt. Einerseits werden die herkömmlichen Testverfahren so modifiziert, dass sie gegenüber einer kontemporären Korrelation der Störgrößen robust sind, vgl. Chang (2002). Ein alternativer Ansatz besteht darin, verallgemeinerte Schätzverfahren (z.B. Verallgemeinerte KQ-Schätzung) zu verwenden, bei denen im ersten Schritt die Korrelationsmatrix aus den Residuen geschätzt wird und im zweiten Schritt eine Transformation vorgenommen wird, mit deren Hilfe eine asymptotisch effiziente Schätzung erreicht werden kann, wie bei Breitung und Das (2005).

Vorliegende Dissertation schlägt neue Methoden vor, die sich explizit mit dieser Kategorie von möglicherweise instationären und kreuz-abhängigen Paneldaten beschäftigen. Typische Anwendungen der hier entwickelten Verfahren sind Mehrländerstudien, bei denen mehrere Länder über ein paar Jahrzehnte beobachtet werden. Die hier implementierten empirischen Anwendungen sind das Testen auf Anwesenheit eines deterministischen Zeittrends in der Investitionsquote am BIP europäischer Volkswirtschaften und die Untersuchung von OECD-Zinssätzen.

Im ersten Kapitel wird ein Paneltest auf Anwesenheit eines Zeittrends vorgeschlagen. Es ist wichtig, einen Test durchführen zu können, der zwischen einem Zeittrend und einem rein zufälligen Verhalten in Gegenwart von Korrelation unterscheiden kann. Die Feststellung der Anwesenheit eines linearen Trends ist von großer praktischer Bedeutung und kann in unterschiedlichen Gebieten angewendet werden, von der technischen Analyse von Finanzmärkten bis zur Vorhersage der globalen mittleren Temperatur. Der vorgeschlagene Test ist in kreuz-abhängigen, heterogenen Paneldaten anwendbar, so lange die Anzahl der Einheiten kleiner als die Anzahl der Zeitbeobachtungen ist. Ferner wird gezeigt, wie der Test angewendet werden kann, wenn Unsicherheit darüber besteht, von welchem Integrationsgrad die stochastischen Komponenten des Modells sind. Dies erfolgt mit der Hilfe des Subsamplingverfahrens. Das Subsampling lässt eine beliebige Anzahl von Einheiten zu und arbeitet besser in kleinen Stichproben als die asymptotische Variante des Tests. Anschließend wird die Methode benutzt, um zu testen, ob der Anteil der Investitionsquote am BIP europäischer Volkswirtschaften einen Trend besitzt. Als Ergebnis wird ein Abwärtstrend auf Panelebene gefunden, wobei die Evidenz für einzelne Länder unschlüssig ist.

In dem nächsten Kapitel wird das Thema des Kointegrationstestens in kreuz-abhängigen Paneldaten behandelt. Ein Einzelgleichungs-Fehlerkorrekturmodell wie das von Banerjee, Dolado and Mestre (1998) eingesetzte Modell wird mit Hilfe nonlinearer Instrumentenschätzung geschätzt und getestet. Diese Art der Schätzung wurde von Chang (2002) für den erweiterten Dickey-Fuller Test vorgeschlagen. Beim Testen der Nullhypothese "keine Kointegration" für eine einzelne Einheit ergibt sich ohne Exogenitätsannahmen eine Teststatistik, die asymptotisch standardnormalverteilt ist. In Panels, die Kreuz-Korrelation oder Kointegration zwischen Einheiten aufweisen, sind hier die einzelnen Teststatistiken asymptotisch unabhängig, was zu einer Panelteststatistik führt, die robust gegen solche Arten von Abhängigkeit ist.

Der dritte Kapitel untersucht in einem ökonometrischen Kontext die einfache Idee, p-Werte einer Reihe statistischer Tests zu kombinieren. Die Idee, Signifikanzniveaus von korrelierten Statistiken zu kombinieren, ist der Biometrie und Medizin entnommen, wo oft nicht die Teststatistiken selber, sondern ihre entsprechenden p-Werte kombiniert wurden. Die Methode ist für Studien geeignet, bei denen die Ergebnisse einzelner Experimente sehr unterschiedlich und deshalb auch unmöglich direkt zu vereinen sind. Die klassische Variante dieses Verfahrens, die Inverse Normale Methode, setzt Unabhängigkeit der einzelnen Teststatistiken voraus, um asymptotische Normalität der gemeinsamen Teststatistik zu erhalten. In diesem Teil der Dissertation wird die von Hartung (1999) vorgeschlagene Korrektur ausführlich behandelt, die für eine bestimmte Korrelationsmatrix der transformierten p-Werte vorgesehen ist. Zuerst wird hier gezeigt, dass die sogenannte Modifizierte Inverse Normale Methode auch für allgemeinere Korrelationsmatrizen gültig ist. Zweitens wird mit Hilfe von Kopulas eine notwendige und ausreichende Bedingung für asymptotische Normalität hergeleitet. Als nächstes werden Anwendungen auf kreuz-abhängige, stationäre wie auch integrierte Paneldaten betrachtet. Dieser Ansatz ist nicht nur für viele Arten von empirischen Fragestellungen adäquat, sondern auch einfach zu implementieren und anzuwenden. Die Monte-Carlo-Simulationen zeigen, dass die Methode gut abschneidet, auch wenn einige der Annahmen nicht erfüllt sind.

Der letzte Kapitel wendet moderne Panelverfahren auf OECD-Zinssätze an, wie auch auf deren Abstände. Genauer gesagt werden hier zum einen Paneleinheitswurzeltests, zum anderen Prozeduren eingesetzt, die einzelne Signifikanzniveaus zu einem gemeinsamen p-Wert kombinieren. Die empirischen Ergebnisse hängen davon ab, ob eine homogene autoregressive Struktur für alle Länder angenommen wird. Die Annahme unabhängiger Einheiten scheint genau so wichtig zu sein. Zugunsten der Stationarität der Zinssätzen und der Zinsabstände spricht lediglich relativ schwache Evidenz, wenn Kreuz-Korrelation mit Hilfe der Modifizierten Inversen Normalen Methode berücksichtigt wird.

Als Gesamtergebnis meiner Dissertation werden bessere Methoden für Paneluntersuchungen bei instationären Daten und Kreuz-Abhängigkeiten geliefert. Mit diesen Methoden werden auch Antworten zu ausgewählten empirischen Fragen gesucht. Zudem lassen sich die Ergebnisse auch zur Untersuchung weiterer ökonomischer Hypothesen einsetzen, sofern diese langfristige Beziehungen implizieren.

Bibliography

Adda, J. and Gonzalo, J. (1996), *P*-Values for Non-Standard Distributions with an Application to the DF Test, *Economics Letters*, Vol. 50, pp. 155-160.

Andrews, D.W.K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817-858.

Bai, J. and S. Ng (2004), A PANIC Attack on Unit Roots and Cointegration, *Econometrica* 72, 1127-1177.

Baltagi, B.H. and Ch. Kao (2000), Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey, in: B.H. Baltagi, (ed.), *Advances in Econometrics,* Vol. 15, Elsevier Science, Oxford.

Banerjee, A. (1999), Panel Data Unit Roots and Cointegration: An Overview, *Oxford Bulletin of Economics and Statistics*, Special Issue, pp. 609-629.

Banerjee, A., J.J. Dolado and R. Mestre (1998), Error-correction Mechanism Tests for Cointegration in a Single-equation Framework, *Journal of Time Series Analysis* 19, 267-283.

Banerjee, A., M. Marcellino and Ch. Osbat (2004), Some Cautions on the Use of Panel Methods for Integrated Series of Macroeconomic Data, *Econometrics Journal* 7, 322-340.

Bickel, P. J. and Doksum, K.A. (2001), *Mathematical Statistics: Basic Ideas and Selected Topics*, 2nd edition, Volume I, Prentice Hall, New York.

Breitung, J. (2000), The Local Power of some Unit Root Tests for Panel Data, in: B.H. Baltagi (ed.), *Advances in Econometrics*, Vol. 15, Elsevier Science, Oxford.

Breitung, J. (2002), Nonparametric Tests for Unit Roots and Cointegration, *Journal of Econometrics* 108, 343-363.

Breitung, J. and S. Das (2005), Panel Unit Root Tests Under Cross Sectional Dependence, *Statistica Neerlandica* 59, 1-20.

Breitung, J. and M. H. Pesaran (2006), Unit Roots and Cointegration in Panels, forthcoming in Matyas, L. and P. Sevestre (ed.), *The Econometrics of Panel Data*, 3^{rd} edition, Kluwer Academic Publishers.

Bunzel, H. and T.J. Vogelsang (2003), Powerful Trend Function Tests That are Robust to Strong Serial Correlation with an Application to the Prebisch-Singer Hypothesis, Working Papers Series, Department of Economics of Iowa State University.

Canjels, E. and M.W. Watson (1997), Estimating Deterministic Trends in the Presence of Serially Correlated Errors, *Review of Economics and Statistics* 19, 271-288.

Carrion-i-Silvestre, J.L. (2003), A Simple Test of Additive Outlier Detection for Stationary Gaussian Time Series, "Analisi Quantitativa Regional" Research Group, University of Barcelona, mimeo.

Chang, Y. (2002), Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency, *Journal of Econometrics* 110, 261-292.

Chang, Y. (2005), Residual-Based Tests for Cointegration in Dependent Panels, presented at the 12th Conference on Panel Data, June 24th-26th, 2005, Copenhagen.

Chang, Y., J.Y. Park and P.C.B. Phillips (2001), Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors, *Econometrics Journal* 4, 1-36.

Chang, Y. and W. Song (2005), Unit Root Tests for Panels in the Presence of Short-run and Long-run Dependencies: Nonlinear IV Approach with Fixed N and Large T, mimeo.

Choi, I. (2001), Unit Root Tests for Panel Data, *Journal of International Money and Finance* 20, 249-272.

Demetrescu, M. (2006), Panel Unit Root Testing with Nonlinear Instruments for $AR(\infty)$ Processes, mimeo.

Demetrescu, M., V. Kuzin and U. Hassler (2005), Long Memory Testing in the Time Domain, forthcoming *Econometric Theory*.

Fisher, R.A. (1954), *Statistical Methods for Research Workers*, 12^{th} edition, Oliver&Boyd, Edinburgh.

Hartung, J. (1999), A Note on Combining Dependent Tests of Significance, *Biometrical Journal* 41, 849-855.

Hassler, U., M. Demetrescu and A.-I. Tarcolea (2005), Asymptotically Normal Unit Root Testing: Extensions to Panel Data and Structural Breaks, presented at the 11th Conference on Panel Data, June 4th-6th, College Station, Texas.

Hatanaka, M. and Yamada, H. (2003), *Co-trending: A Statistical System Analysis of Economic Trends*, Springer, Tokyo.

Heike, H.-D. and Târcolea, C. (2000), *Grundlagen der Statistik und Wahrscheinlichkeitsrechung*, Oldenbourg.

Herrndorf, N. (1984), A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables, *Annals of Probability* 12, 141-153.

Heston, A., R. Summers and B. Aten, (2002), Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002.

Im, K.S. and M.H. Pesaran (2003), On the Panel Unit Root Tests using Nonlinear Instrumental Variables, mimeo.

Im, K.S., M.H. Pesaran and Y. Shin (2003), Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics* 115, 53-74.

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

de Jong, R.M. and C.-H. Wang (2005), Further Results on the Asymptotics for Nonlinear Transformations of Integrated Time Series, *Econometric Theory* 21, 413-430.

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

Kasahara, Y. and S. Kotani (1979), On the Limit Process for a Class of Additive Functionals of Recurrent Diffusion Processes, *Zeitschrift für Wahrscheinlichkeitstheorie verw. Gebiete* 49, 133-153.

Kiefer, N.M., T.J. Vogelsang and H. Bunzel (2000), Simple Robust Testing of Regression Hypothesis, *Econometrica* 68, 695-714.

Kim, T.-H., S. Pfaffenzeller, T. Rayner and P. Newbold (2003), Testing for Linear Trend with Application to Relative Primary Commodities Prices, *Journal of Time Series Analysis* 24, 539-551.

Kirchgässner, G. and J. Wolters (1993), Does the DM Dominate the Euro Market? An Empirical Investigation, *The Review of Ecomics and Statistics*, 75, pp. 773-778.

Kirchgässner, G. and J. Wolters (1995), Interest Rate Linkages in Europe Before and After the Introduction of the European Monetary System, *Empirical Economics*, 20, pp. 435-454.

Kuzin, V. (2005), Recursive Demeaning and Deterministic Seasonality, *Statistics & Probability Letters* 72, 195-204.

Kwiatkowski, D, P.C.B. Phillips, P. Schmidt and Y. Shin (1992), Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, *Journal of Econometrics* 54, 159-178.

Levin, A., C.-F. Lin and C.-S.J. Chu (2002), Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties, *Journal of Econometrics* 108, 1-24.

Levine, R. and D. Renelt (1992), A Sensitivity Analysis of Cross-Country Growth Regressions, *The American Economic Review* 82, 942-963.

Liptak, T. (1958), On the combination of independent tests, *Magyar Tudomanyos Akademia Matematikai Kutato Intezetenek Kozlemenyei*, Vol. 3, pp. 171-197.

Lo, A.W. (1991), Long-Term Memory in Stock Market Prices, *Econometrica* 59, 1279-1313.

Lobato, I.N. (2001), Testing that a Dependent Process is Uncorrelated, *Journal of the American Statistical Association* 96, 1066-1076.

Lütkepohl, H. (1996), *Handbook of matrices*, Wiley and Sons, Chichester.

MacKinnon, J.G. (1996), Numerical Distribution Functions for Unit Root and Cointegration Tests, *Journal of Applied Econometrics* 11, pp. 601-618.

Maddala, G.S. and S. Wu (1999), A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, Special Issue, pp. 631-652.

Madsen, J.B. (2002), The Causality Between Investment and Economic Growth, *Economics Letters* 74, 157-163.

McCoskey, S. and C. Kao (1998), A Residual-Based Test of the Null of Cointegration in Panel Data, *Econometric Reviews* 17, 57-84.

Moon, H.R. and B. Perron (2004), Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics* 122, pp. 81-126.

Nelsen, R.B. (1998), *An Introduction to Copulas*, Springer-Verlag, New York.

Newey, W.K. and K.D.West (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

O'Connell, P.G.J. (1998), The Overvaluation of Purchasing Power Parity, *Journal of International Economics* 44, 1-19.

Park, J.Y. and P.C.B. Phillips (1988), Statistical Inference in Regressions with Integrated Processes: Part1, *Econometric Theory* 4, 468-498.

Park, J.Y. and P.C.B. Phillips (1999), Asymptotics for Nonlinear Transformations of Integrated Time Series, *Econometric Theory* 15, 269-298.

Park, J.Y. and P.C.B. Phillips (2001), Nonlinear Regressions with Integrated Time Series, *Econometrica* 69, 117-162.

Pedroni, P. (2004), Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis, *Econometric Theory* 20, 597-625.

Politis, D.N. and J.P. Romano (1994), Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions, *Annals of Statistics* 22, 2031-2050.

Politis, D.N., J.P. Romano and M. Wolf (1999), *Subsampling*, Springer-Verlag, New York.

Robinson, P.M. (1994), Efficient Tests of Nonstationary Hypotheses, *Journal of the American Statistical Association* Vol. 89, pp. 1420-1437.

Shin, Y. (1994), A Residual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration, *Econometric Theory* 10, 91-115.

Sun, H. and S.G. Pantula (1999), Testing for Trends in Correlated Data, *Statistics and Probability Letters* 41, 87-95.

Tanaka, K. (1999), The Nonstationary Fractional Unit Root, *Econometric Theory* Vol. 15, pp. 549-582.

Van Zwet, W.R. and Oosterhoff, J. (1967), On the Combination of Independent Test Statistic, *The Annals of Mathematical Statistics*, Vol. 38, pp. 659-680.

Vogelsang, T.J. (1998), Trend Function Hypothesis Testing in the Presence of Serial Correlation, *Econometrica* 66, 123-148.

Wolters, J. (2002), Uncovered Interest Rate Parity and the Expectation Hypothesis of the Term Structure: Empirical Results for the U.S. and Europe, in: I. Klein and S. Mittnik (eds.), *Contributions to Modern Econometrics: From Data Analysis to Economic Policy. In Honour of Gerd Hansen*, Kluwer, pp. 271-282.

Curriculum Vitae

Adina Ioana Tarcolea
geboren am 12. Januar 1979 in Bukarest

Wissenschaftlicher Werdegang

06/1997	Abitur in Bukarest
1997-2001	Studium an der Universität Bukarest, Fachbereich Mathematik
06/2001	Studienabschluss (Diplom-Mathematikerin)
2001 - 2002	Wissenschaftliche Mitarbeiterin an der Technischen Universität Bukarest (Fachgebiet Statistik, Prof. Dr. Mariana Craiu)
2002 - 2003	DAAD-Stipendium an der Technischen Universität Darmstadt (Fachgebiet Statistik und Ökonometrie, Prof. Dr. Hans-Dieter Heike)
seit 2003	Wissenschaftliche Mitarbeiterin an der Goethe-Universität Frankfurt am Main (Institut für Statistik und Methoden der Ökonometrie, Prof. Dr. Uwe Hassler)

Ehrenwörtliche Erklärung

Ich habe die vorgelegte Dissertation selbst verfaßt und dabei nur die von mir angegebenen Quellen und Hilfsmitteln benutzt. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, sowie alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Frankfurt am Main, den 6. Juni 2006