Testing for Cointegration in Panel Data with Cross-Sectional Dependence

Antonia Arsova

Testing for Cointegration in Panel Data with Cross-Sectional Dependence

Von der Fakultät Wirtschaftswissenschaften
der Leuphana Universität Lüneburg zur Erlangung des Grades

Doktorin der Wirtschafts- und Sozialwissenschaften
- Dr. rer. pol. -

genehmigte Dissertation von
Antonia Arsova

geboren am 26.09.1985
in Sofia, Bulgarien

Eingereicht am: 15.02.2019

Mündliche Verteidigung (Disputation) am: 17.05.2019

Erstbetreuerin und -gutachterin: Apl. Prof. Dr. Deniz Karaman Örsal
Zweitgutachter: Prof. Dr. Jürgen Jacobs
Drittgutachter: Prof. Dr. Jörg Breitung

Die einzelnen Beiträge des kumulativen Dissertationsvorhabens sind wie folgt veröffentlicht:

Arsova, A., and Örsal, D. D. K., (2018). Likelihood-based panel cointegration test in the presence of a linear time trend and cross-sectional dependence. *Econometric Reviews* 37:10, 1033-1050, https://doi.org/10.1080/07474938.2016.1183070.

Arsova, A., and Örsal, D. D. K., (2019). Intersection tests for the cointegrating rank in dependent panel data. *Communications in Statistics – Simulation and Computation*, https://doi.org/10.1080/03610918.2018.1489552.

Veröffentlichungsjahr: 2019

Acknowledgements

The creation of this thesis has been a long journey which wouldn't have come to a successful end without the help and support of many people and institutions, whom I wish to express my gratitude.
First of all, I would like to thank my supervisor Adj. Prof. Deniz Örsal for accepting me as her PhD student and for carefully guiding me through my endeavours. I have always highly appreciated her kind support, her readiness to discuss anything at any time and the freedom she granted me to explore my own ideas.
I am also thankful to the German Research Foundation (DFG) for their financial support through the projects KA-3145/1-1 and KA-3145/1-2.
My gratitude goes also to Prof. Maroussia Bojkova and Prof. Racho Denchev for their encouragement in the very beginning, and to Plamen Trayanov and Tsvetomira Tsenova for helpful and inspiring discussions.
Beginning and completing my PhD studies would have been impossible without the inspiration, help and encouragement of Boris Blagov and Dimitar Kodjabachev. Thank you! I also want to thank all my friends who have been there for me during the ups and downs of my doctoral path and have encouraged me to keep going.
Last but not least, I wish to thank my family. My wholehearted gratitude goes to my mother, my father, my brother and my aunt for their unconditional love and support which empower me to pursue my dreams.

To my mother

Abstract

This cumulative thesis extends the econometric literature on testing for cointegration in nonstationary panel data with cross-sectional dependence. Its self-contained chapters consist of two publications and two publication manuscripts which present three new panel tests for the cointegrating rank and an empirical study of the exchange rate pass-through to import prices in Europe.

The first chapter introduces a new cointegrating rank test for panel data where the dependence is assumed to be driven by unobserved common factors. The common factors are first estimated and subtracted from the observations. Then an existing likelihood-ratio panel cointegration test is applied to the defactored data. The distribution of the test statistic, computed from defactored data, is shown to be asymptotically equivalent to that of a test statistic computed from cross-sectionally independent data.

The second chapter proposes a new panel cointegrating rank test based on a multiple testing procedure, which is robust to positive dependence between the individual units' test statistics. The assumption of a certain type of positive dependence is shown by simulations not to be violated in panels with dependence structures commonly assumed in practice. The new test is applied to find empirical support of the monetary exchange rate model in a panel of eight OECD countries.

The third chapter puts forward a new panel cointegration test allowing for both cross-sectional dependence and structural breaks. It employs known individual likelihood-ratio test statistics accounting for breaks in the deterministic trend and combines their p-values by a novel modification of the Inverse Normal method. The average correlation between the probits is inferred from the average cross-sectional correlation between the residuals of the individual VAR models in first differences.

The fourth chapter studies the exchange rate pass-through to import prices in a panel of nineteen European countries through the prism of panel cointegration. Empirical evidence supporting a theoretical long-run equilibrium relationship between the model's variables is found by the newly proposed panel cointegration tests. Two different panel regression models, which take both cointegration and cross-sectional dependence into account, provide most recent estimates of the exchange rate pass-through elasticities.

Contents

1

Introduction and overview

1.1 Introduction

In a seminal paper "Spurious regressions in econometrics" Granger and Newbold (1974) demonstrated how regressions between wholly unrelated nonstationary variables give rise to statistically significant estimates. They called this phenomenon a "spurious" regression, and as a remedy to the problem recommended to estimate such models after taking first differences of the variables – at least "until a really satisfactory procedure is available". Although the authors acknowledged that it might not be a universally applicable solution, estimating econometric models in first differences rather than in levels enjoys popularity among practitioners even today. One problem with this modelling approach, although not immediately recognized, is that it mainly focuses on dynamics in the short run and hence potential links between the variables in the long run may get lost.

Continuing work on this subject by investigating nonstationary processes and how they interact in the long run, Granger published "Some properties of time series data and their use in econometric model specification" in 1981. In it the author coined the term "co-integration" to describe the co-movements of two or more nonstationary, or integrated, stochastic processes, which do not drift too far apart in the long run, but remain tied closely together.

Not before long another very influential article came along: "Co-integration and error-correction: representation, estimation and testing" by Engle and Granger (1987). They showed how cointegrated processes can be modelled in an error-correction framework and proposed new methods for estimation and hypothesis testing, essentially providing the necessary "satisfactory procedure" sought after more than a decade earlier. With the availability of this new statistical toolbox cointegrated regression models quickly became the primary workhorse for testing and estimation of long-run relationships between economic variables. In recognition of the importance of this work Robert Engle and Sir Clive Granger were awarded with the Nobel prize for economics in 2003.

The nonstationary behaviour of a time series is usually described by its order of integration. Integrated variables of order d (denoted as $I(d)$) are such that they become covariance-stationary after taking repeated differences d times. By the definition given by Engle and Granger (1987), integrated variables of order d are said to be cointegrated if there exists one or more nontrivial linear combination(s) of them which has a lower order of integration. Over time, many economic variables have been found to exhibit behaviour consistent with the properties of $I(1)$, also termed "unit-root"[1], processes. Therefore in practice, unless noted otherwise, cointegration most commonly refers to $I(1)$ variables sharing common unit-root stochastic trend(s), thus having a stationary (or $I(0)$) long-run relationship.

Testing for the integration and cointegration order of economic variables established itself as an essential pre-modelling step in the late 1980s, aided by the surge of newly proposed tests for detecting unit roots and cointegration. Cointegration tests may be viewed as multivariate extentions of unit root tests, and two main types thereof can be distinguished: residual-based and system tests. Residual-based cointegration tests, on the one hand, look at regression models of integrated variables and examine whether the regression error series contains unit roots. They are therefore suitable to detect the presence of a single cointegrating relationship provided the regression model is correctly specified. System tests, on the other hand, may be used to detect not only the presence, but also the number of the equilibrium relationships among several variables, which is known as the *cointegrating rank*. They are also more flexible as they do not depend on the choice of a variable for the normalization of the cointegrating relationship. One very prominent system test still widely employed today is the likelihood-based cointegrating rank test of Johansen (1988). It became the base for many further developments and extensions, some of which are pursuit in this thesis.

In the early 1990s it was recognized that the power of popular univariate unit root tests could be rather low, especially for short time series; this critique naturally applies also to cointegration tests. However, as extending the data in the time dimension (T) is not always possible or desirable because of issues with data availability or quality, a way of

[1]The term "unit root" describes one of the roots of the characteristic equation of an $I(1)$ process, which lies on the unit circle, while the rest lie within.

improving power had to be found elsewhere. This led to the emergence of *panel* unit root and cointegration tests, where the amount of data is increased in the cross-sectional dimension (N). Such extention is deemed meaningful on the belief that time series from different units, e.g. countries, industries, or firms, exhibit similar dynamics over time. The combined information set is therefore supposed to be more revealing for the stochastic properties of their driving forces. An extensive review of the early panel cointegration tests can be found in Örsal (2009).

A common trait characterizing the so-called *first generation* of tests is the assumption of independence between the cross-sectional units. This assumption significantly simplifies the derivation of the asymptotic distributions of the proposed panel test statistics, in most cases by invoking the Central Limit Theorem for a panel statistic computed as the standardized average of the individual statistics. It was nevertheless quickly recognized that albeit convenient, the assumption of independence is almost never fulfilled in practice. The consequences of unattended cross-sectional dependence could be quite detrimental for both estimation and hypothesis testing. For example, the least squares estimator may become biased, inconsistent or inefficient. The presence of dependence also invalidates the limiting distributions of panel unit root and cointegration test statistics assuming independence, so that the probability of a Type I error, a.k.a. the size of a test, may become much larger than the nominal significance level.

Cross-sectional dependence arises naturally in macroeconomic panel data as a result of common shocks, spatial spillover effects and general interdependence due to the increasingly tighter economic and financial links between countries. Various approaches to tackle the issues it brings about in the existing panel unit root and cointegration tests began to emerge in the early 2000s. These range from estimating the variance-covariance matrix of the stacked panel model residuals directly by generalized least squares as in Seemingly Unrelated Regression Equations (SURE), through eliminating the dependence by constructing suitable instrumental variables, to bootstrapping for approximating the limiting distribution of panel test statistics. More coherent treatment of this subject can be found in the excellent review by Breitung and Pesaran (2008). Two approaches that have become particularly popular in the development of such *second-generation* panel unit root and cointegration tests that are robust to cross-sectional dependence merit a separate mention. The first one assumes that the dependence stems from unobserved common factors which affect the different units with heterogeneous loadings. The common factors may then be either estimated from the data, so that decomposition of the observed variables into unobserved common and idiosyncratic components takes place, or they may be approximated by functions of the observables, e.g. by cross-sectional averages. The second approach, in contrast, does not assume a particular dependence structure in the data generating processes (DGP) of the time series, but rather seeks to combine the significance levels of the dependent individual statistics into a panel statistic. Such meta-analytic tests are naturally more flexible as they permit much more heterogeneity in the DGP. Both of the aforementioned approaches have been advanced in this thesis to propose new panel cointegration tests. Recently, a *third generation* of panel unit root and cointegration tests has been on the rise: such that allow for both cross-sectional dependence and structural breaks in the data generating process. Issues posed by structural breaks have become particularly relevant in the years following the Global Financial Crisis, as it has caused a break in many, if not all, economic and financial time series. Unattended structural breaks, however, may generally cause a stationary series seem nonstationary and hence distort the results of unit root and cointegration tests, similarly to what unattended cross-sectional dependence does to first-generation panel tests. Hence new tools capable of handling them both simultaneously have been called for.

1.2 Overview

This thesis, which consists of two publications and two publication manuscripts, extends the literature on panel cointegration testing on both theoretical and empirical level. First, it develops two new second-generation tests and one third-generation test for the cointegrating rank in panel data. Second, it contributes to empirical economics by applying these new tests to investigate the presence of long-run equilibrium in different economic systems whose existence is predicted by theory.

A common trait shared by the proposed panel cointegration tests is that they are all based on existing likelihood-based tests for the cointegrating rank which carry out detrending of the observed data prior to the computation of the test statistics. The indvidual-unit test underlying both new second-generation panel tests is the one developed by Saikkonen and Lütkepohl (2000) (henceforth SL test). They demonstrate that estimating the deterministic linear trend by generalized least squares (GLS) and subtracting it from the observations leads to improved power properties of the test in comparison with Johansen's (1988) test allowing for a linear time trend. The same GLS-detrending procedure is followed also by Trenkler, Saikkonen and Lütkepohl (2007), whose test allowing for structural breaks serves as the basis for the newly proposed third-generation panel cointegration test. What differs between the newly developed tests is *a)* the way the cross-sectional dependence is handled; and *b)* the construction of the panel test statistic. In the first panel test, presented in Chapter 2, the assumed source of cross-sectional dependence is eliminated asymptotically from the observed data. Therefore, the panel test statistic, computed as the standardized average of the individual-unit test statistics, has a standard normal limiting distribution by the Central Limit Theorem for independent and identically distributed observations. The other type of second-generation test, presented in Chapter 3, and the third-generation test, presented in Chapter 4, look rather into incorporating the cross-sectional dependence into the computation of the panel test statistic. Hence they both resort to p-value combination methods which are robust to cross-sectional dependence.

In terms of empirical contributions the thesis presents results of testing for cointegration by the newly proposed tests using recent data in three different areas. First, a long-run relationship between the nominal exchange rate and monetary fundamentals, postulated by the monetary exchange rate model, is revealed by means of the test proposed in Chapter 3. Second, the theoretical equilibrium between house prices and personal income is examined in Chapter 4 by the third-generation panel test in view of the structural break caused by the Global Financial Crisis. Finally, an empirical model for the exchange rate pass-through, which looks at the long-run linkage between import prices and exchange rate movements, is investigated in Chapter 5. The existence of a cointegrating relationship is established by the tests proposed in this thesis. It serves as the basis for the estimation of two different cointegrated panel regression models whose results are then compared. A brief summary of each chapter, highlighting both the theoretical and the empirical contributions, is presented next.

Chapter 2 presents the publication of Arsova and Örsal (2018), which develops a new panel cointegrating rank test allowing for a linear time trend and cross-sectional dependence. It combines individual SL test statistics as the panel SL test by Örsal and Droge (2014) and may be viewed as an extention of the latter test to dependent panels. The dependence is assumed to be driven by unobserved common factors, which are extracted from the data by the method of principal components. Thus the observed time series are essentially decomposed into unobserved common and idiosyncratic components, allowing their integrating and cointegrating properties to be determined separately. Testing for cointegrating rank zero is then performed by applying the panel SL test to the defactored data. Testing for higher cointegrating ranks, on the other hand, is carried out by applying the panel SL test to estimates of the hypothesized stochastic trends. Three theorems lead to the proof that, under certain conditions, the asymptotic distribution of the panel test statistic, derived from defactored data, is the same as that of the statistic computed from the cross-sectionally independent idiosyncratic components. In other words, any bias arising from estimating the unobserved common factors and their loadings, is shown to be asymptotically negligible as the dimensions of the panel T and N grow indefinitely. A Monte Carlo simulation study demonstrates that the proposed testing procedure has good finite sample properties, which are preferable than those of an alternative test, developed earlier for an equivalent setting. The framework of Arsova and Örsal (2018) is extended further by Örsal and Arsova (2017) to propose new panel rank tests. The latter are based on methods for combination of p-values and are shown to exhibit even better finite sample properties in some situations.

Chapter 3 presents the forthcoming publication of Arsova and Örsal (2019), which takes a different approach to dealing with the cross-sectional dependence in the proposed testing procedure for the panel cointegrating rank. Two new panel tests, which differ in the underlying individual likelihood-based cointegration statistics, emerge from this procedure. The first test combines the individual likelihood-based cointegration statistics of Johansen (1988) (henceforth J), while the second one is again based on the SL statistics. These panel tests too assume that the dependence may stem from unobserved common factors, but they rather focus on the cointegrating rank of the observed variables directly, without carrying out decomposition into unobserved components. In order to allow for both cross-sectional dependence and as much heterogeneity in the different units as possible, a p-value combination method – the intersection test of Simes (1986) – is employed for the computation of the panel test statistic. Initially developed for multiple testing, the intersection test presents an improvement of Bonferroni's procedure, and in earlier literature it has been shown to observe the nominal significance level when the individual test statistics are characterized by a certain type of positive dependence. In practice it is difficult to verify whether this particular type of positive dependence, namely *multivariate totally positive of order 2* (MTP_2), arises between the individual J or SL cointegrating rank statistics in the panel. We therefore propose to empirically measure whether such dependence is likely to appear in dependent panel data driven by unobserved common factors by resorting to simulation methods. For this aim a multivariate version of Kendall's tau is adapted as a measure for MTP_2, and in a simulation study this measure is computed by means of the empirical copula of the individual statistics. The results confirm that the MTP_2 assumption, necessary for Simes' procedure, is not violated in panel data with this commonly assumed dependence structure. The finite-sample properties of the Simes-J and Simes-SL panel cointegrating rank tests, examined in the Monte Carlo study, prove to be satisfactory with the power increasing significantly as the cross-sectional dimension grows. The newly proposed tests are then applied to investigate the validity of the monetary exchange rate model for a panel of eight OECD countries. The presence of a cointegrating relation, as predicted by economic theory, is established.

Chapter 4 presents a publication manuscript, co-authored with Adj. Prof. Deniz Örsal, which proposes a panel test for the cointegrating rank allowing for both cross-sectional dependence and structural breaks. The underlying individual rank statistics are those of Trenkler et al. (2007) (henceforth TSL), whose test allows for up to two breaks with known locations in the intercept and slope of the deterministic linear trend in the DGP. The p-values of the TSL test statistics are combined by means of a novel modification of the Inverse Normal method, which aims at making it robust to cross-sectional dependence. The panel test statistic of the Inverse Normal method is computed as the standardized average of the probits of the individual p-values, which in the case of independent one-sided statistics has a standard normal distribution. Assuming that the cross-sectional dependence can adequately be captured by the average correlation between the probits, a modification for dependent statistics has been put forward by Hartung (1999). It, however, relies on an estimator of the correlation between the probits, which is based on the single $(N \times 1)$ vector of observations on the probits that is available in practice and is thus not very precise, and on somewhat arbitrary variance inflation factor in the computation of the panel test statistic. Our contribution lies in proposing a novel, more reliable estimator of the correlation between the probits, which is inferred from a quantity easily measurable in practice. This quantity is the average absolute cross-sectional correlation between the innovations to the individual DGPs in the panel, which is consistently estimated from the residuals of the individual vector autoregressive (VAR) models. The functional relationship between the correlation between the innovations to the individual DGPs and the correlation between the probits is approximated by a response surface regression on observations generated in a large-scale simulation study. The finite-sample properties of the new correlation-augmented inverse normal (CAIN) test are shown to be preferable to those of the Hartung's (1999) modification. The CAIN-TSL test is also compared to another panel cointegration test allowing for structural breaks and is shown to be superior. As an empirical illustration it is applied to investigate whether a long-run equilibrium between house prices and personal income exists in the US housing market using most recent data. Despite modelling the outset of the Global Financial Crisis as a structural break in the trend of the DGP, we fail to establish cointegration on the panel level. This leads us to the conclusion that house prices have deviated from their theoretical driving force – personal income – in the period prior to the Global Financial Crisis. Equilibrium, however, seems to be restored in the years after the crisis as indicated by the results of the Simes-SL test, which establishes cointegration at the panel level for the period 2008-2018.

Finally, Chapter 5 presents a publication manuscript which takes a cointegration approach to estimating the exchange rate pass-through into import prices in a panel of nineteen European countries. Despite the fact that a long-run equilibrium between import prices, nominal exchange rate and other macroeconomic determinants of import prices is predicted by economic theory, in recent empirical studies the presence of cointegration has either been overlooked, or it could not be established. Therefore, the first contribution of this study is to investigate the existence of cointegration by means of the newly proposed tests. A single cointegrating relationship between the observed import prices, nominal exchange rate, domestic demand and a proxy for producers' costs is found by the Simes-SL test. To gain insight into the driving forces of this relationship, decomposition of each observed series into unobserved common and idiosyncratic components is carried out. Both types of components turn out to be non-stationary and cointegrated, whereas the cointegration among the observed variables is found to be driven mainly by two global stochastic trends. The second contribution of this study is to present recent estimates of the short- and long-run pass-through elasticities, taking cointegration into account. For this aim two different cointegrated panel regression models, both of them allowing for cross-sectional dependence, are estimated. Despite the differences in the estimation techniques, the results of the two models are qualitatively and quantitatively similar and indicate incomplete low pass-through at the panel level both in the short- and in the long-run.

Bibliography

Arsova, A., Örsal, D.D.K., 2018. Likelihood-based panel cointegration test in the presence of a linear time trend and cross-sectional dependence. Econometric Reviews 37, 1033–1050.

Arsova, A., Örsal, D.D.K., 2019. Intersection tests for the cointegrating rank in dependent panel data. Forthcoming in Communications in Statistics: Simulation and Computation , doi:10.1080/03610918.2018.1489552.

Breitung, J., Pesaran, M.H., 2008. Unit Roots and Cointegration in Panels, in: Mátyás, L., Sevestre, P. (Eds.), The Econometrics of Panel Data. Springer Berlin Heidelberg, Berlin, Heidelberg. volume 46 of *Advanced Studies in Theoretical and Applied Econometrics*, pp. 279–322.

Engle, R.F., Granger, C.W.J., 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica 55, 251.

Granger, C., 1981. Some properties of time series data and their use in econometric model specification. Journal of Econometrics 16, 121–130.

Granger, C., Newbold, P., 1974. Spurious regressions in econometrics. Journal of Econometrics 2, 111–120.

Hartung, J., 1999. A Note on Combining Dependent Tests of Significance. Biometrical Journal 41, 849–855.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 12, 231–254.

Örsal, D.D.K., 2009. Essays on panel cointegration testing. Ph.D. thesis. Humboldt-Universität zu Berlin.

Örsal, D.D.K., Arsova, A., 2017. Meta-analytic cointegrating rank tests for dependent panels. Econometrics and Statistics 2, 61–72.

Örsal, D.D.K., Droge, B., 2014. Panel cointegration testing in the presence of a time trend. Computational Statistics & Data Analysis 76, 377–390.

Saikkonen, P., Lütkepohl, H., 2000. Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process. Journal of Time Series Analysis 21, 435–456.

Simes, R.J., 1986. An improved Bonferroni procedure for multiple tests of significance. Biometrika 73, 751–754.

Trenkler, C., Saikkonen, P., Lütkepohl, H., 2007. Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break. Journal of Time Series Analysis 29, 331–358.

2

Likelihood-based panel cointegration test in the presence of a linear time trend and cross-sectional dependence

Antonia Arsova, Deniz Dilan Karaman Örsal

This article proposes a new likelihood-based panel cointegration rank test which extends the test of Örsal and Droge (2014, *Computational Statistics and Data Analysis* 76: 377–390) (henceforth panel SL test) to dependent panels. The dependence is modelled by unobserved common factors which affect the variables in each cross-section through heterogeneous loadings. The data are defactored following the panel analysis of nonstationarity in idiosyncratic and common components (PANIC) approach of Bai and Ng (2004, *Econometrica* 72(4): 1127–1177) and the cointegrating rank of the defactored data is then tested by the panel SL test. A Monte Carlo study demonstrates that the proposed testing procedure has reasonable size and power properties in finite samples.

Keywords: Common factors; cross-sectional dependence; likelihood-ratio; panel cointegration rank test; time trend.
JEL classification: C12, C15, C33

Earlier versions of this paper were presented at the 13th IWH-CIREQ Macroeconometric Workshop in Halle and at the Conference on Cross-Sectional Dependence in Panel Data Models in Cambridge. The authors thank workshop and conference participants and four anonymous referees for many helpful comments and suggestions. Financial support by the German Research Foundation (DFG) through the project KA-3145/1-1 is also gratefully acknowledged.

This chapter is derived in full from an article published in *Econometric Reviews* 2018, vol. 37, No. 10, pp. 1033–1050, copyright Taylor & Francis, available online: http://www.tandfonline.com/10.1080/07474938.2016.1183070.

3

Intersection tests for the cointegrating rank in dependent panel data

Antonia Arsova, Deniz Dilan Karaman Örsal

This paper takes a multiple testing perspective on the problem of determining the cointegrating rank in macroeconomic panel data with cross-sectional dependence. The testing procedure for a common rank among the panel units is based on Simes' (1986, *Biometrika* 73(3): 751–754) intersection test and requires only the p-values of suitable individual test statistics. A Monte Carlo study demonstrates that these simple tests are robust to cross-sectional dependence and have reasonable size and power properties. A multivariate version of Kendall's tau is used to test an important assumption underlying Simes' procedure for dependent statistics. The proposed method is illustrated by an empirical application.

Keywords: Common factors; cross-sectional dependence; likelihood-ratio; multiple testing; panel cointegration rank test.
JEL classification: C12, C15, C33

Financial support by the German Research Foundation (DFG) through the project KA-3145/1-1 is gratefully acknowledged. The data for the empirical analysis was in part collected during a research stay of the second author at the Humboldt Universität in Berlin. The first author would like to thank Plamen Trayanov for helpful discussions. The authors are also grateful to two anonymous referees for their constructive comments and suggestions.

This chapter is derived in full from an article published in *Communications in Statistics – Simulation and Computation* 2019, copyright Taylor & Francis, available online: http://www.tandfonline.com/10.1080/03610918.2018.1489552.

4

A panel cointegrating rank test with structural breaks and cross-sectional dependence

Antonia Arsova, Deniz Dilan Karaman Örsal

This paper proposes a new panel cointegrating rank test which allows for a linear time trend with breaks and cross-sectional dependence. The new correlation-augmented inverse normal (CAIN) test is based on a novel modification of the inverse normal method and combines the p-values of individual likelihood-ratio trace statistics. A Monte Carlo study demonstrates its robustness to cross-sectional dependence and its superior size and power properties compared to other meta-analytic tests used in practice. The test is applied to investigate the long-run relationship between regional house prices and personal income in the United States in view of the structural break introduced by the Global Financial Crisis.

Keywords: Panel cointegrating rank test, structural breaks, cross-sectional dependence, common factors, likelihood-ratio, time trend

JEL classification: C12, C15, C33

Financial support by the German Research Foundation (DFG) through the project KA-3145/1 – 2 is gratefully acknowledged.

4.1 Introduction

Panel unit root and cointegration tests have been developed since the early 2000s with the aim to increase the power of single-unit tests. The so-called "first generation" tests rely on the assumption of independence between the panel units. In macroeconomic panel data, however, cross-sectional dependence arises naturally due to common shocks or spillover effects. If not accounted for, it may bias the outcome of the tests by inflating the type-I error rate above the nominal significance level. Another issue often observed in longer time series is that of structural breaks; it may also invalidate the test results if left unattended.

Focusing on cointegrating rank testing, we address both issues simultaneously by extending the cointegrating rank test of Trenkler et al. (2007) (henceforth the TSL test) to panel data with cross-sectional dependence. Under the assumptions of the test structural breaks are allowed in the deterministic parts of the data generating process (DGP), i.e. level and trend slope, but not in the cointegrating vector. This is in line with its interpretation as a long-run equilibrium relationship between the variables in the system. Furthermore, this framework allows for structural breaks both under the null and the alternative hypothesis, as the breaks do not affect the stochastic properties of the DGP. Our preference for the TSL test over another likelihood-based alternative, the cointegrating rank test of Johansen et al. (2000) (henceforth JMN test), is motivated by its superior finite-sample properties demonstrated by Trenkler et al. (2007) and our own simulations.

Panel cointegration testing in the presence of structural breaks and cross-sectional dependence has only recently gained attention from researchers, leading to the development of the so called "third-generation tests". Westerlund and Edgerton (2008), for example, propose a simple panel test for no cointegration allowing for a level shift and a break in the cointegrating relation, but not for a break in the deterministic trend. They assume that the cross-sectional dependence is driven by stationary unobserved common factors, which might be seen as restrictive in practice. Banerjee and Carrion-i Silvestre (2015) relax these restrictions in their no-cointegration test and propose test with level shifts, level shifts and break in the cointegrating relation, or level and trend shifts. However, in the latter case they allow only for homogeneous number of breaks and break dates across the panel units. They point out that "the difficulty [in allowing for heterogeneous break dates] essentially lies in the dependence of the critical values of the tests on the location of the break dates when trend breaks are present". While the same holds true for the TSL test, our approach to combining information from individual cross-sections into a panel test accommodates heterogeneous number of breaks and break dates across units.

We extend the TSL test to the panel setting resorting to a new p-value combination method which allows the p-values to be correlated. p-value combination approaches offer much more flexibility than traditional pooling of individual test statistics, as they allow the specification of the deterministic terms, the lag order, the number and the location of the breaks, and even the time span of the data to vary over cross-sections.

Recent research on panel unit root and cointegration testing has benefited significantly from the "reinvention" of already existing methods for combining possibly dependent p-values. One example is a modification of the Bonferroni procedure proposed by Simes (1986) and employed by Hanck (2013) in his panel unit root test. In a Monte Carlo study he demonstrates its robustness to cross-sectional dependence induced by common factors. This avenue is further explored in the direction of cointegrating rank testing in dependent panels by Arsova and Örsal (2019), who also show empirically that a sufficient condition for the validity of Simes' procedure is not violated in a common-factor-driven panel framework.[1]

In the present work we adopt a new, augmented version of yet another p-value combination method initially proposed by Stouffer et al. (1949) — the inverse normal test. The novelty of our approach lies in that we explicitly model the degree of cross-sectional correlation between the probits of the individual statistics and use it as a variance-inflation factor in the panel test statistic. In this regard the proposed test is similar to the modified inverse normal method of Hartung (1999), which also uses an estimate of the cross-sectional correlation of the probits, however in combination with a somewhat arbitrary correction factor κ.[2] Our estimator for the unobserved correlation of the probits is modelled as a function of a quantity easily measurable in practice, namely the average absolute cross-sectional correlation of the residuals of the individual VAR models. In a Monte Carlo study we demonstrate that our correlation-augmented inverse normal (CAIN) method for combining p-values of individual TSL tests has good size and power properties in finite samples. Its performance is preferable to that of Hartung's method over the whole range of possible values for the cross-sectional correlation. The test can be used to determine the cointegrating rank of the observed time series even when the cross-sectional dependence is driven by unobserved common factors, without decomposing the data into idiosyncratic and common components and testing these separately.

The remainder of the paper is organized as follows. The next section briefly reviews the test of Trenkler et al. (2007), while its extension to panel data by a modification of the inverse normal method is given in Section 3. Section 4 presents the response surface approach to modelling the cross-sectional correlation between the probits of

[1] Sarkar (1998) shows that for the Simes' procedure to hold for p-values of dependent statistics, their multivariate distribution has to be *multivariate totally positive* (MTP_2). We refer to the latter article for a definition of this property.

[2] Hartung's approach has been used to account for cross-sectional dependence in panel unit root tests by Demetrescu et al. (2006) and Costantini and Lupi (2013).

the individual test statistics. Section 5 discusses the results of a Monte Carlo study comparing the CAIN-TSL test with other meta-analytic approaches and with the test of Westerlund (2006). Section 6 illustrates the use of the CAIN-TSL test by investigating the equilibrium relationship between real house prices and real personal income in the USA, and the last section concludes. Supplementary material is provided in the Appendix.

4.2 The TSL test for the cointegrating rank

Our aim is to develop a panel test for the cointegrating rank which allows for structural breaks in the deterministic parts of the DGP. A natural first step is to take an existing single-unit test and to extend it to the panel setting. Two well-known alternatives are the likelihood-based test of Johansen et al. (2000) and its GLS-detrended counterpart proposed by Trenkler et al. (2007). Both tests allow for up to two breaks in the level and/or trend slope, whereas the break dates are assumed to be known.[3]

Trenkler et al. (2007) find that the JMN test with a single trend break displays excessive size distortions in systems of larger dimensions while their test is correctly sized. Our own Monte Carlo simulations confirm these findings also when two breaks in the trend slope are present; the results are briefly summarized in Table 4.1.

Table 4.1: Empirical size at 5% nominal level of the JMN and TSL tests with two structural breaks, $H_0 : r = 0$

Test	$T = 50$	$T = 100$	$T = 200$	$T = 500$
$\text{JMN}_{0.25,0.5}$	0.545	0.209	0.116	0.070
$\text{JMN}_{0.25,0.75}$	0.531	0.206	0.111	0.075
$\text{TSL}_{0.25,0.5}$	0.047	0.046	0.042	0.044
$\text{TSL}_{0.25,0.75}$	0.052	0.051	0.043	0.045

Subscripts denote break locations at relative sample lengths.
Simulations based on a 3-variate VAR(2) process as in Wagner and Hlouskova (2010) with 5000 replications.

The JMN test is severely oversized for small T, with the size distortions persisting even for $T = 500$. The TSL test, in contrast, has approximately correct size at the 5% level for all values of T. Its preferable finite-sample properties are therefore our main motivation for extending the TSL test to the panel setting. Next we briefly describe its assumptions and the computation of the test statistic.

Let the observed data $Y_{it} = (Y_{1,it}, \ldots, Y_{m,it})'$ for cross-sectional unit i ($i = 1, \ldots, N$) be generated by a stochastic VAR(s_i) process X_{it} added to a deterministic process. The

[3]Theoretically more than two breaks could be specified, however response surface regressions approximating the limiting distributions of the test statistics are available for a maximum of two breaks for both tests. Including more than two breaks may also be seen as problematic in time series of a limited time span.

latter consists of a constant, linear time trend and structural breaks in both the level and the trend slope at known individual-specific time(s) τ_i:

$$Y_{it} = \mu_{0i} + \mu_{1i}t + \delta_{0i}d_{it} + \delta_{1i}b_{it} + X_{it}, \quad t = 1, \ldots, T_i. \tag{4.1}$$

Here μ_{ji} and δ_{ji} $(j = 0, 1)$ are unknown $(m \times 1)$ parameter vectors, while d_{it} and b_{it} are dummy variables defined by $d_{it} = b_{it} = 0$ for $t < \tau_i$, and $d_{it} = 1$ and $b_{it} = t - \tau_i + 1$ for $t \geq \tau_i$. We note that both the break dates and the number of breaks are assumed to be known. The number of breaks (one or two) is allowed to vary across units. The break dates are assumed to occur at individual-specific fixed fractions of the sample size: $\tau_i = [T_i\eta_i]$ with $0 < \underline{\eta_i} < \eta_i < \overline{\eta_i}$, where $\underline{\eta_i}$ and $\overline{\eta_i}$ are specified real numbers and $[\cdot]$ denotes the integer part of the argument. In other words, the breaks are assumed not to occur in the very beginning or in the very end of the sample, while $\underline{\eta_i}$ and $\overline{\eta_i}$ are allowed to be arbitrarily close to 0 and 1, respectively. The stochastic processes X_{it} are assumed to be at most $I(1)$ and cointegrated with cointegrating rank r_i, $0 \leq r_i \leq m - 1$:

$$X_{it} = A_{1i}X_{i,t-1} + \ldots + A_{s_i i}X_{i,t-s_i} + \varepsilon_{it}, \quad t = 1, \ldots, T_i. \tag{4.2}$$

It is assumed that the $(m \times 1)$ vector ε_{it} is distributed as $i.i.d.(0, \Omega_i)$, where Ω_i is a positive definite matrix for each i. Further it is assumed that ε_{it} have finite moments of order $(4 + \nu)$ for some $\nu > 0$, $\forall i$.

Denoting the pairwise cross-sectional correlations of the elements of ε_{it} by $\rho_{il,jk} := \text{corr}(\varepsilon_{it,l}, \varepsilon_{jt,k})$ for $i, j = 1, \ldots, N$ and $l, k = 1, \ldots, m$, we make the following assumptions.

Assumption 1 The average absolute pairwise cross-sectional correlation between the innovations to the same variable converges to some fixed value $\rho_\varepsilon > 0$ for all t as $N \to \infty$:

$$\lim_{N\to\infty} \frac{1}{mN(N-1)} \sum_{i\neq j}^{N} \sum_{l=1}^{m} |\rho_{il,jl}| = \rho_\varepsilon \,. \tag{4.3}$$

Assumption 2 The average absolute pairwise cross-sectional correlation between the innovations to different variables converges to zero for all t as $N \to \infty$:

$$\lim_{N\to\infty} \frac{1}{mN(N-1)} \sum_{i\neq j}^{N} \sum_{l\neq k}^{m} |\rho_{il,jk}| = 0. \tag{4.4}$$

We note that Assumption 1 is not strictly necessary, but rather eases the interpretation of the estimated average absolute correlation coefficient $\hat{\rho}_\varepsilon$. Albeit seemingly restrictive, Assumption 2 is rather a technical one, necessary for the computation of a suitable estimator of ρ_ε. It is motivated by the fact that strong correlations are less likely to occur between the innovations to different variables across units compared to those between the same variables. Furthermore, our Monte Carlo simulations demonstrate that the proposed panel test is robust to a certain degree of deviation from this assumption, hence it should not hinder the applicability of the test in practice. Both assumptions hold when, for example, a spatial type of dependence is assumed, or when the dependence is driven by variable-specific common factors. As such we describe unobserved shocks which affect the same variable over the cross-sections, but do not or only marginally affect other variables.

The computation of the individual TSL test statistics proceeds by estimating the deterministic terms by reduced rank regression taking into account the structural breaks, and then computing a likelihood-ratio (LR) trace statistic from the trend-adjusted observations. For details on the procedure we refer to Trenkler et al. (2007).

4.3 The correlation-augmented inverse normal test

Let p_i denote the p-values of the individual TSL statistics for units $i = 1, \dots, N$. Let t_i denote the corresponding probits, i.e. $t_i = \Phi^{-1}(p_i)$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. Assuming independence of the individual test statistics (and hence of their p-values and the corresponding probits t_i), the inverse normal test has a standard $N(0,1)$ limiting distribution:

$$t = \frac{\sum_{i=1}^{N} \Phi^{-1}(p_i)}{\sqrt{N}} = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N}} \Rightarrow N(0,1). \tag{4.5}$$

The inverse normal method was first introduced to the panel unit root testing literature by Choi (2001), who demonstrates by simulations that under cross-sectional independence it outperforms other p-value combination alternatives, in particular Fisher's inverse Chi-square method employed for panel unit root testing by Maddala and Wu (1999). As a further advantage of the inverse normal test Choi (2001) points out its applicability to panels with both finite and infinite cross-sectional dimension N.

Assuming multivariate normal distribution of the probits, Hartung (1999) proposes a modification of the weighted inverse normal method to accommodate dependence between the original test statistics. The real-valued weights of the individual probits are denoted by λ_i and are such that $\sum_{i=1}^{N} \lambda_i \neq 0$. In practice it is often assumed that $\lambda_i = 1, \forall i$. The dependence is captured by a single correlation coefficient ρ_t, which can be interpreted as a "mean correlation approximating the case of possibly different correlations between the transformed statistics" (Hartung, 1999). The variance of the denominator in (4.5) is then augmented with an estimator of the correlation between the individual probits $\hat{\rho}_{\mathrm{t}}^{*}$:

$$t\left(\hat{\rho}_{\mathrm{t}}^{*},\kappa\right)=\frac{\sum_{i=1}^{N}\lambda_i t_i}{\sqrt{\sum_{i=1}^{N}\lambda_i^2+\left[\left(\sum_{i=1}^{N}\lambda_i\right)^2-\sum_{i=1}^{N}\lambda_i^2\right]\left[\hat{\rho}_{\mathrm{t}}^{*}+\kappa\cdot\sqrt{\frac{2}{(N+1)}}(1-\hat{\rho}_{\mathrm{t}}^{*})\right]}}. \tag{4.6}$$

The estimator $\hat{\rho}_{\mathrm{t}}^{*}$ is computed as

$$\hat{\rho}_{\mathrm{t}}^{*}=\max\{-\frac{1}{N-1},\hat{\rho}_{\mathrm{t}}\},\ \text{where} \tag{4.7}$$

$$\hat{\rho}_{\mathrm{t}}=1-\frac{1}{N-1}\sum_{i=1}^{N}\left(t_i-\frac{1}{N}\sum_{i=1}^{N}t_i\right)^2. \tag{4.8}$$

The correction term $\kappa\sqrt{\frac{2}{(N+1)}}(1-\hat{\rho}_{\mathrm{t}}^{*})$, which simply scales the standard deviation of $\hat{\rho}_{\mathrm{t}}$ by a factor κ, aims to avoid a systematic underestimation of the denominator in eq. (4.6). For the κ parameter Hartung suggests two alternative values: $\kappa_1=0.2$ and $\kappa_2=0.1\cdot\left(1+\frac{1}{N-1}-\hat{\rho}_{\mathrm{t}}^{*}\right)$, where κ_2 is suitable mainly for smaller $\hat{\rho}_{\mathrm{t}}^{*}$. However, he provides no guidance as to where the threshold between "small" and "large" $\hat{\rho}_{\mathrm{t}}^{*}$ should be.

Demetrescu et al. (2006) are the first to employ Hartung's modified inverse normal method to develop a panel unit root test allowing for cross-sectional dependence by generalizing Hartung's approach in two directions. First, they prove that the correlation between the individual test statistics needs not be constant for the limiting $N(0,1)$ distribution to hold. Second, they show that a necessary and sufficient condition for the limiting normality of the panel test statistic is the multivariate distribution of the individual test statistics to have a Gaussian copula. This theoretical result is, however, difficult to verify in practice; hence the authors proceed to demonstrate by simulation the applicability of Hartung's method to dependent individual ADF unit root tests. Employing the version of the test with $\kappa_1=0.2$ and unit weights $\lambda_i=1,\forall i$, they find that for medium and strong cross-sectional correlation it generally observes the nominal significance level at 5% and 10%. For weak correlation the test is rather undersized and therefore not recommended for use with more than 5 cross-sectional units.

Following the approach of Demetrescu et al. (2006), Costantini and Lupi (2013) use Hartung's modification with the value $\kappa_1=0.2$ in their simple panel-CADF test for unit roots. In order to mitigate the size distortions for weak cross-sectional dependence, they propose a type of "switching algorithm" between the regular and the modified inverse normal test ((4.5) and (4.6), respectively) based on the outcome of the CD test for cross-sectional correlation of Pesaran (2015). When the CD test rejects the null of weak correlation, (4.6) is used, otherwise Choi's test (4.5) is utilized. We argue that such strategy is sub-optimal because of two reasons. Firstly, the implicit null hypothesis and the finite-sample properties of the CD test depend on the relative expansion rates of T and N. It tends to over-reject if T is large relative to N and the exponent of cross-sectional dependence $\alpha \in (1/4, 1/2]$, a situation which in our view might well occur in macroeconometric panels.[4] Hence Hartung's modification might get preferred over Choi's test too often, leading to a loss of power. Secondly, the CD statistic accounts for the *average* correlation coefficient so that in its calculation large correlations of the opposite sign will cancel out. We argue that even negative correlations between the innovations to the DGPs lead to positive correlation between the individual test statistics. As our Monte Carlo study demonstrates, a better way to quantify the degree of cross-sectional dependence is to look at the mean *absolute* residual correlation.

We therefore propose a new, improved version of the modified inverse normal test for combination of correlated individual TSL test statistics for cointegration with structural breaks. This new test, which we name CAIN-TSL test, is based on Hartung's (1999) test statistic (4.6) with unit weights, $\lambda_i = 1, \forall i$. It differs from it in that it employs a novel empirical estimate $\tilde{\rho}_t$ of the average correlation between the probits ρ_t. In order to accommodate a certain degree of heterogeneity in the correlation coefficients $\rho_{t_{i,j}} := \mathrm{corr}(t_i, t_j)$, we make the same assumptions as Demetrescu et al. (2006).

Assumption 3

$$\lim_{N\to\infty} \frac{1}{N(N-1)} \sum\sum_{i\neq j} \rho_{t_{i,j}} = \rho_t, \quad \rho_t \in (0,1), \text{ and} \tag{4.9}$$

$$\lim_{N\to\infty} \frac{1}{N(N-1)} \sum\sum_{i\neq j} (\rho_{t_{i,j}} - \rho_t)^2 = 0. \tag{4.10}$$

Assumption 4 The individual TSL test statistics have a Gaussian copula.

Assumption 3 allows the correlation matrix of the probits to gradually approach a constant correlation matrix as $N \to \infty$, which is a natural consequence of the assumed convergence of the pairwise innovation correlations in Assumption 1. Assumption 4, as shown by Demetrescu et al. (2006), is a necessary and sufficient condition for the limiting distribution of the inverse normal test to be standard normal. Whether this assumption is met in practice for individual unit root or cointegration statistics, is not known; we leave such investigation for future research. However, as our Monte Carlo study demonstrates, in the presence of cross-sectional dependence the proposed CAIN test performs much better than the standard inverse normal method without correlation augmentation, despite the possibility that Assumption 4 might not hold. Hence in practice one would still be better off by correcting for existing cross-sectional dependence rather than ignoring it.

The CAIN panel test statistic for the composite null hypothesis $H_0 : r_i = r$, $\forall i =$

[4] We refer to Pesaran (2015) for a definition of the exponent of cross-sectional dependence.

$1, \dots, N$ against the alternative $H_1 : r_i > r$ for at least one i is given by

$$t(\tilde{\rho}) = \frac{\sum_{i=1}^{N} t_i}{\sqrt{N + (N^2 - N) \cdot \tilde{\rho}_{\mathrm{t}}}}. \tag{4.11}$$

Conjecture 1 *Provided that Assumptions 1–4 hold, the limiting distribution of the test statistic $t(\tilde{\rho})$ under the null hypothesis is approximate standard normal.*

The proposed estimator $\tilde{\rho}_{\mathrm{t}}$ is based on an empirical estimate of the link between the cross-sectional correlation of the innovations to the individual VAR processes and the correlation between the probits of the individual TSL statistics. Since there is no analytic expression for it, no proof of the conjecture can be provided in the strict mathematical sense. As noted by Hartung (1999), the statistic converges to the $N(0,1)$ distribution when the estimator of the correlation between the probits is consistent. Therefore, in the end of Section 4 we outline the arguments by which $\tilde{\rho}_{\mathrm{t}}$ is expected to be consistent.

4.4 Response surface regressions for the correlation of the individual probits

Common shocks to the innovations of the DGPs for the individual units, or generally cross-sectional correlation of the innovations leads to dependence between the individual TSL test statistics. This dependence transfers also to the p-values and their probits, and therefore has to be taken into account in the construction of the inverse normal test statistic in order to achieve a correctly sized test.

We argue that ρ_{t} can be inferred from its origin ρ_{ε}, which, in turn, can be consistently estimated in practice as N and T grow. It is, however, difficult to derive analytically how correlation between the innovations translates into correlation between the LR statistics, as the latter are complex non-linear functions of the observed data. We therefore resort to simulation methods to estimate this link. In our large-scale simulation study we estimate the average correlation of the probits $\bar{\bar{\rho}}_{\mathrm{t}}$ for different values of the absolute cross-sectional correlation between the innovations ρ_{ε}, controlling for the system dimension m, the hypothesized cointegrating rank r, and the time and cross-sectional dimensions T and N, respectively. We then estimate the relationship $\bar{\bar{\rho}}_{\mathrm{t}} = g(\rho_{\varepsilon}, m, r)$ by response surface regression. As in practice ρ_{ε} is unobserved, we replace it by a consistent estimate $\hat{\rho}_{\varepsilon}$ to compute $\tilde{\rho}_{\mathrm{t}} = g(\hat{\rho}_{\varepsilon}, m, r)$ to use in the test statistic (4.11).

Simulations design We estimate the relationship between ρ_{ε} and ρ_{t} in a large-scale simulation study. The data are generated according to (4.1) with the stochastic processes X_{it} following an m-variate VAR(1) Toda processes (see e.g. Toda, 1995) for

$m = 2, 3, 4, 5$:

$$X_{it} = \begin{pmatrix} I_d & 0_{(d\times r)} \\ 0_{(r\times d)} & \Psi_r \end{pmatrix} X_{i,t-1} + \varepsilon_{it}, \tag{4.12}$$

where I_d denotes the identity matrix of dimension $d = m - r$, $0_{(d\times r)}$ is a zero matrix of the corresponding dimension and r is the cointegrating rank of the process. Ψ_r is a diagonal matrix of dimension $(r \times r)$ with diagonal elements $\psi_i \sim i.i.d.\, U(0, 1)$, $i = 1, \ldots, r$ which govern the dynamics of the stationary elements of the process for $r > 0$; the uniform $(0, 1)$ distribution of ψ_i provides for generality. As the individual LR trace statistics of the TSL test are invariant to the values of the deterministic terms μ_{0i}, μ_{1i}, δ_{0i} and δ_{1i} for given VAR order s_i and break date(s) τ_i (Trenkler et al., 2007, p. 340), we have set $\mu_{ji} = \delta_{ji} = 0$, $j = 0, 1$.

For the estimation of the individual LR trace statistics for each $i = 1, \ldots, N$ the number of breaks is randomly chosen between 1 and 2 with 50% chance each, and the break locations are chosen at relative sample length(s) $\lambda_i \sim i.i.d.\, U(0.15, 0.85)$. In the case of two breaks the minimal distance between them is set to $0.2T$. These values are in line with what is usually assumed in the literature on cointegration with structural breaks – the breaks are not allowed to be too close to the beginning or to the end of the sample, neither are they allowed to be too close to each other.

The innovations ε_{it} are drawn from a multivariate normal distribution with variance-covariance matrix Σ, where $\Sigma = R \otimes \Omega$ is generated as in Wagner and Hlouskova (2010) with

$$R = \begin{pmatrix} 1 & \rho_\varepsilon & \cdots & \rho_\varepsilon \\ \rho_\varepsilon & 1 & \cdots & \ddots \\ \vdots & \ddots & \ddots & \rho_\varepsilon \\ \rho_\varepsilon & \cdots & \rho_\varepsilon & 1 \end{pmatrix}_{(N\times N)}, \tag{4.13}$$

and $\Omega_{(m\times m)}$ being a random correlation matrix generated independently for each replication as described in Costantini and Lupi (2013, p. 283).

For each value of m, testing $H_0 : r_i = r_0$ for $r_0 = 0, \ldots, m - 1$ is considered in a separate experiment.

We let ρ_ε vary over a grid of 24 equally spaced values in the range $[0.04, 0.96]$, in order to be able to adequately fit an interpolating curve to the estimated average values $\bar{\hat{\rho}}_{\mathrm{t}}$ of ρ_{t} over $\hat{\rho}_\varepsilon$. Each simulation experiment is repeated 100000 times for three combinations of the panel dimensions T and N: $(T, N) \in \{(500, 5), (500, 10), (1000, 5)\}$.

Estimation of the average correlation of the probits $\bar{\hat{\rho}}_{\mathrm{t}}$ from simulated data

For each combination (T, N) an average $\bar{\hat{\rho}}_{\mathrm{t}}$ is computed by means of Fisher's Z-transformation from the $(N \times N)$ correlation matrix of the probits based on the

$(100000 \times N)$ matrix of independent observations on $(t_1, \ldots, t_N)$.

For given ρ_ε and T, the estimated $\bar{\tilde{\rho}}_{\mathrm{t}}$ is practically invariant to the number of cross-sections. This fact is not surprising, as it is rational to expect that the correlation between the individual probits would depend on the degree of dependence between the processes of any two cross-sections, but not on the number of units. Preliminary simulations corroborate this conjecture – see the right panel of Figure 4.3 in the Appendix. For robustness, the estimated $\bar{\tilde{\rho}}_{\mathrm{t}}$'s for all three combinations of (T, N) are subsequently modelled in the response surface regressions for $\bar{\tilde{\rho}}_{\mathrm{t}} = g(\rho_\varepsilon, m, r)$. In practice, by the Law of Large Numbers, increasing N would lead to more precise estimation of $\hat{\rho}_\varepsilon$ and subsequently also of $\tilde{\rho}_{\mathrm{t}}$. Because of the large number of replications in the current simulation exercise, considering a wider range of values for N is not necessary and we concentrate only on values which are typical for macroeconometric studies.

With regard to the choice of values for the parameter T, we motivate it by the fact that the estimated $\bar{\tilde{\rho}}_{\mathrm{t}}$ converges to ρ_{t} from below as $T \to \infty$ (see the left panel of Figure 4.3), with the differences between $T = 500$ and $T = 1000$ being virtually negligible for all practical purposes. Indeed, for small T the estimated $\tilde{\rho}_{\mathrm{t}}$ might be lower than its asymptotic large-T value and could thus potentially get overestimated. We argue, however, that the use of the large-T $\tilde{\rho}_{\mathrm{t}}$ even for small T's in the panel setting would have beneficial rather than detrimental effects. It is well known that individual likelihood-based cointegration tests tend to be oversized for $H_0 : r = 0$ when T is small, and these size distortions can get magnified in the panel setting as the cross-sectional dimension increases (see, e.g., Demetrescu and Hanck, 2012 and Arsova and Örsal, 2018). Using the asymptotic $\tilde{\rho}_{\mathrm{t}}$ might help mitigate this issue, as slight overestimation of the cross-sectional correlation of the probits might inflate the variance of the panel statistic thus offsetting the inherent size distortions of the individual tests.

Response surface regressions We now turn our attention to modelling the relationship $\bar{\tilde{\rho}}_{\mathrm{t}} = g(\rho_\varepsilon, m, r)$. To obtain the response surface g, we regress $\bar{\tilde{\rho}}_{\mathrm{t}}$ on polynomials of the system dimension m, the cointegrating rank under the null hypothesis r, and the mean absolute correlation between the innovations ρ_ε. No constant is included in the regression and all regressors are multiples of ρ_ε, so that all estimates $\bar{\tilde{\rho}}_{\mathrm{t}}$ of ρ_{t} are equal to 0 when $\rho_\varepsilon = 0$.

The goodness of fit measure of the estimated regression is $R^2 = 0.9993$, which renders the approximation very good for practical purposes. All estimated coefficients, which are significant at the 5%-level, are listed in Table 4.2. They can be used to compute the estimate $\tilde{\rho}_{\mathrm{t}}$ of the unknown correlation of the probits given m, r and the estimated $\hat{\rho}_\varepsilon$ for any system dimension $m \leq 5$. The CAIN-TSL test statistic is then computed as in (4.11) with $\tilde{\rho}_{\mathrm{t}}$ in the place of the unknown ρ_t.

Table 4.2: Coefficients of the response surface $\tilde{\rho}_{\mathrm{t}} = g(\rho_\varepsilon, m, r)$ for the correlation between the probits of the individual TSL statistics

Term	Estimated coefficient	Std. Err.	t-value	$P > \|t\|$	95% Confidence interval
${\rho_\varepsilon}^2$	0.6319575	0.0357830	17.66	0	$(0.561742, 0.702173)$
		0.0309826	20.4	0	$(0.571162, 0.692753)$
$\sqrt{m} \cdot {\rho_\varepsilon}^2$	−0.5193669	0.0251431	−20.66	0	$(-0.568704, -0.470030)$
		0.0211871	−24.51	0	$(-0.560941, -0.477792)$
$\sqrt{m} \cdot {\rho_\varepsilon}^4$	0.2721753	0.0032355	84.12	0	$(0.265826, 0.278524)$
		0.0033330	81.66	0	$(0.265635, 0.278716)$
$\frac{r}{m} \cdot {\rho_\varepsilon}^2$	0.1821374	0.0307753	5.92	0	$(0.121749, 0.242526)$
		0.0260690	6.99	0	$(0.130983, 0.233292)$
$\frac{r}{m} \cdot {\rho_\varepsilon}^4$	−0.0856903	0.0322216	−2.66	0.008	$(-0.148917, -0.022463)$
		0.0293583	−2.92	0.004	$(-0.143299, -0.028082)$
$(r \cdot \rho_\varepsilon)^2$	0.0041125	0.0008434	4.88	0	$(0.002458, 0.005768)$
		0.0007581	5.43	0	$(0.002625, 0.005600)$
$r \cdot {\rho_\varepsilon}^2$	0.0766267	0.0076098	10.07	0	$(0.061694, 0.091559)$
		0.0070263	10.91	0	$(0.062839, 0.090414)$
$r \cdot {\rho_\varepsilon}^4$	−0.1008678	0.0057871	−17.43	0	$(-0.112224, -0.089512)$
		0.0060007	−16.81	0	$(-0.112643, -0.089093)$
$\sqrt{m-r} \cdot {\rho_\varepsilon}^2$	0.1874919	0.0348379	5.38	0	$(0.119131, 0.255853)$
		0.0276625	6.78	0	$(0.133211, 0.241773)$
$\frac{1}{m-r} \cdot {\rho_\varepsilon}^2$	0.1410229	0.0168932	8.35	0	$(0.107874, 0.174172)$
		0.0150575	9.37	0	$(0.111476, 0.170570)$
$\frac{1}{m-r} \cdot {\rho_\varepsilon}^4$	−0.2029126	0.0126466	−16.04	0	$(-0.227728, -0.178097)$
		0.0120002	−16.91	0	$(-0.226460, -0.179365)$
$(m-r)^2 \cdot {\rho_\varepsilon}^2$	0.0052557	0.0009492	5.54	0	$(0.003393, 0.007118)$
		0.0008073	6.51	0	$(0.003672, 0.006840)$
$(m-r)^4 \cdot {\rho_\varepsilon}^4$	−0.0000327	0.0000167	−1.96	0.050	$(-0.000065, 0.000000)$
		0.0000179	−1.83	0.068	$(-0.000068, 0.000002)$

The second row for each term presents robust standard errors and the statistics computed therewith.

Estimation of the average absolute cross-sectional correlation between the process innovations The cornerstone of our proposed solution $\tilde{\rho}_{\mathrm{t}} = g(\rho_\varepsilon, m, r)$ is a consistent estimator $\hat{\rho}_\varepsilon$ of the average absolute cross-sectional correlation of the process innovations. In this regard we follow Pesaran (2015) and estimate $\hat{\rho}_\varepsilon$ from the residuals of the individual VAR(s_i) models under the null hypothesis $H_0 : r = 0$. Our estimation methodology, however, differs from his one in two aspects. First, as we are dealing with a panel of multivariate systems, the average cross-sectional correlation coefficient needs to be redefined for this richer data structure. We assume that strong cross-correlations are more likely to appear between the shocks to the same variables than between those to different variables. Hence averaging has to be performed only across the pairwise correlations between the same-variable residuals in order to avoid underestimation of ρ_ε and subsequently of ρ_{t}. Such underestimation would lead to an oversized panel test. That is, for $\hat{\rho}_{il,jl}$ denoting the estimated sample correlation between the residuals for variable l in units i and j, $\hat{\rho}_\varepsilon$ is to be estimated from the average of $|\hat{\rho}_{il,jl}|$ over $l = 1, \ldots, m$ and $i \neq j$ where $i, j = 1, \ldots, N$. Second, we take the absolute value of the estimated correlations $\hat{\rho}_{il,jl}$. In simple averaging positive and negative correlations will cancel out, thus leading to underestimation of the true degree of cross-sectional correlation.

Therefore, the estimated absolute residual correlations $|\hat{\rho}_{il,jl}|$ are averaged over l and $i \neq j$, resulting in $\hat{\rho}_\varepsilon$:

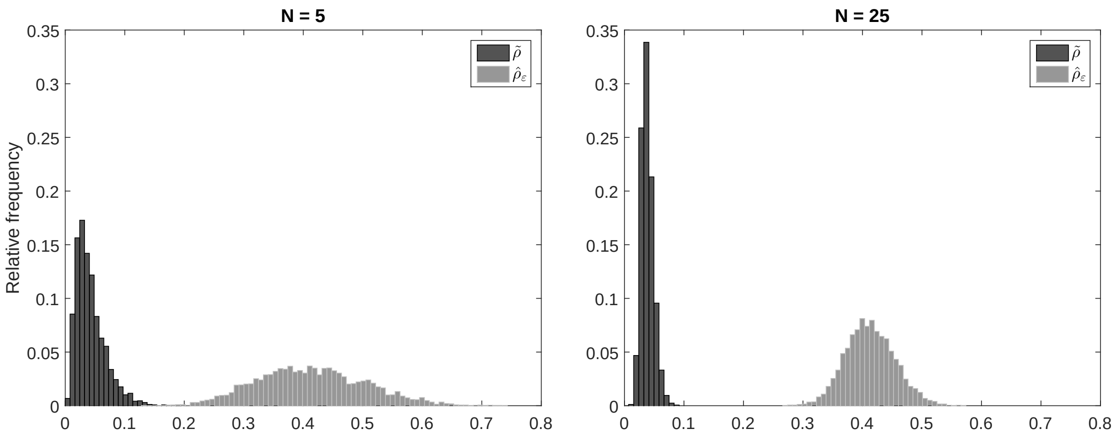
$$\hat{\rho}_\varepsilon = \frac{2}{mN(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{l=1}^{m} |\hat{\rho}_{il,jl}|. \tag{4.14}$$

Estimating $\hat{\rho}_\varepsilon$ only under $H_0 : r_i = 0, \forall i$ in the beginning of the sequential testing procedure for $r = 0, \ldots, m-1$ requires further justification. Ideally, the correlation between the probits would be inferred from the correlation between the residuals under each null hypothesis and then used in the correlation-augmented inverse normal panel test. In order to keep the procedure feasible for practical work, however, we decide to do this inference only once by determining $\hat{\rho}_\varepsilon$ only under the null of no cointegration, similarly to the lag orders s_i of the VAR processes. We note that the estimation error in $\hat{\rho}_\varepsilon$ is negligible under the true $H_0 : r_i = 0, \forall i, m$, while $\hat{\rho}_\varepsilon$ gets only slightly underestimated for higher true cointegrating ranks (see Table 4.6 in the Appendix). The resulting slight underestimation of ρ_{t}, however, would be negligible in practice for moderate ρ_ε; for higher ρ_ε it could even be beneficial, given that the TSL test is severely undersized when testing for higher ranks under the null.

Consistency of the proposed estimator For the approximate $N(0,1)$ limiting distribution of the CAIN-TSL test statistic to hold, the estimator $\tilde{\rho}_{\mathrm{t}}$ must be consistent. As discussed, we propose to estimate $\tilde{\rho}_t$ as a continuous function g of the dimension of the system m, the hypothesized cointegrating rank r and the mean absolute correlation ρ_ε between the innovations of the individual DGPs. The function g results from an ordinary least squares (OLS) regression of mean values $\bar{\hat{\rho}}_{\mathrm{t}}$, estimated from simulated data, onto r, m and ρ_ε. The simulations design reflects the assumptions made and allows for as much heterogeneity in the data-generating processes over the different cross-sections as possible. The estimated values of ρ_{t} are then averaged by means of Fisher's Z-transformation to mitigate any potential bias. Therefore, by the properties of the OLS estimator, $\tilde{\rho}_{\mathrm{t}} = g(r, m, \rho_\varepsilon)$ is unbiased and consistent for ρ_{t}. In practice, however, ρ_ε is unobserved, and has to be estimated from the data. Provided that the residuals from the individual VAR models are homoscedastic and serially uncorrelated, $\hat{\rho}_\varepsilon$ can be consistently inferred from them as in (4.14); consistency follows by the Law of Large Numbers as $N \to \infty$ treating m as fixed. Therefore consistency of $\tilde{\rho}_t$ follows from the consistency of the estimator $\hat{\rho}_\varepsilon$ and the Continuous Mapping Theorem for the mapping g.

An illustration based on simulation evidence is provided in Figure 4.1. It is clear that increasing the cross-sectional dimension of the panel leads to more concentrated distributions of both $\hat{\rho}_\varepsilon$ and $\tilde{\rho}_{\mathrm{t}}$.

Figure 4.1: Empirical distributions of $\hat{\rho}_\varepsilon$ and $\tilde{\rho}_{\mathrm{t}}$ under $H_0 : r_i = 0, \forall i$, 3-variate VAR(2) process with a multifactor-error structure and a diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$, $T = 100$, 5000 replications



Step-by-step outline of the testing procedure

In order to illustrate the simplicity of the proposed testing approach, we summarize it with the following five steps.

1. Compute the TSL statistic and its corresponding p-value as outlined in Trenkler et al. (2007) under $H_0 : r = 0$ for each unit. As a by-product, save the residuals from the estimated individual VAR models.

2. Estimate the $(Nm \times Nm)$ sample correlation matrix of the residuals. If the time span of the series varies over the cross-sections, estimation of the correlation matrix would be based on the balanced panel restricted by the shortest time series.
   Compute $\hat{\rho}_\varepsilon$ as in (4.14).

3. Estimate $\tilde{\rho}_{\mathrm{t}} = g(\hat{\rho}_\varepsilon, m, r)$ using the response surface coefficients in Table 4.2.

4. Compute the panel test statistic by combining the probits of the individual p-values as in (4.11).

5. The CAIN-TSL test can be applied for testing $H_0 : r_i = r, \forall i$ at each step $r = 0, \ldots, m - 1$ of the sequential rank testing procedure. If the composite null hypothesis $H_0 : r_i = 0$, $\forall i$ is rejected, repeat steps 1 and 4, without re-estimating $\hat{\rho}_\varepsilon$. As $\tilde{\rho}_{\mathrm{t}}$ depends on the cointegrating rank r, it needs to be re-estimated.

4.5 Monte Carlo study

4.5.1 Simulation study design

The finite sample properties of the CAIN-TSL are first examined by simulations in an empirically relevant case using a three-variate VAR(2) DGP as in the study of Wagner and Hlouskova (2010). The test is next compared to the panel cointegration test of Westerlund (2006), which as well allows for structural breaks in the deterministic terms.

The general form of the DGP of Wagner and Hlouskova (2010) is:

$$Y_{it} = \mu_{0i} + \mu_{1i}t + \delta_{0i}d_{it} + \delta_{1i}b_{it} + X_{it}, \tag{4.15}$$

$$X_{it} = \begin{pmatrix} a^i_{11} & 0 & 0 \\ 0 & a^i_{12} & 0 \\ 0 & 0 & a^i_{13} \end{pmatrix} X_{i,t-1} + \begin{pmatrix} a^i_{21} & 0 & 0 \\ 0 & a^i_{22} & 0 \\ 0 & 0 & a^i_{23} \end{pmatrix} X_{i,t-2} + u_{it}, \tag{4.16}$$

$$u_{it} = \gamma_i' f_t + \varepsilon_{it}, \tag{4.17}$$

$$\varepsilon_{it} \sim i.i.d.\, N(0, \Omega_i). \tag{4.18}$$

The cointegrating properties of the process are determined by the roots q^i_{1j}, q^i_{2j} of the autoregressive polynomial, which are linked to the coefficients of the autoregressive matrices by $a^i_{1j} = \frac{1}{q^i_{1j}} + \frac{1}{q^i_{2j}}$ and $a^i_{2j} = -\frac{1}{q^i_{1j}q^i_{2j}}$, $j = 1, 2, 3$.

Following Wagner and Hlouskova (2010), for a system with cointegrating rank zero we set $q^i_{1j} = 1$ and $q^i_{2j} \sim U(1.8, 3)$, $j = 1, 2, 3$. Power is investigated in a setting when all roots are sufficiently away from unity (case A) and also in a near-unit root setting (case B). For cointegrating rank one in case A we let $q^i_{11} \sim U(1.3, 1.7)$ or, in case B, $q^i_{11} \sim U(1, 1.3)$, while $q^i_{21} \sim U(1.5, 2.5)$, $q^i_{1j} = 1$ and $q^i_{2j} \sim U(1.8, 3)$ for $j = 2, 3$ in both cases. Finally, for cointegrating rank two we again let $q^i_{11} \sim U(1.3, 1.7)$ for case A or $q^i_{11} \sim U(1, 1.3)$ for case B. The remaining roots for both cases are $q^i_{12}, q^i_{2j} \sim U(1.5, 2.5)$ for $j = 1, 2$, while $q^i_{13} = 1$ and $q^i_{23} \sim U(1.8, 3)$. All roots are drawn separately for each unit.

We let a 3-dimensional vector of variable-specific common factors $f_t \sim i.i.d.\, N(0, I_3)$ drive the cross-sectional dependence through heterogeneous loadings. The factor loadings γ_i are simulated as diagonal $(k \times m)$ matrices with (a) $i.i.d.\, U(-0.4, 0.4)$, (b) $i.i.d.\, U(0, 1)$ or (c) $i.i.d.\, U(-1, 3)$ entries, drawn separately for each unit. The robustness of the CAIN-TSL test against violations of Assumption 2 is investigated by letting γ_i be an unrestricted matrix with $i.i.d.\, U(0, 1)$ or $i.i.d.\, U(-1, 3)$ elements. This corresponds to all factors affecting all variables simultaneously.

The TSL test statistic is invariant to the actual values of the deterministic terms as long as the number and the location of the breaks are correctly specified in its estimation.

Therefore, throughout we set $\mu_{ji} = \delta_{ji} = 0$, $j = 0, 1$ as in Trenkler et al. (2007). We again allow for a random number of breaks (1 or 2, with 50% chance each) and the break locations are chosen at relative fraction(s) of the sample size $\lambda_i \sim i.i.d.\, U(0.15, 0.85)$. In the case of two breaks the minimal distance between them is set to $0.2T$. The processes X_{it} are initialised with 0 and the first 50 observations are discarded to mitigate the effect of initial values. The individual-specific random correlation matrices Ω_i of the idiosyncratic errors ε_{it} are simulated, as before, as described in Costantini and Lupi (2013).

We consider all combinations of $T \in \{100, 200\}$ and $N \in \{5, 15, 25\}$. The lag order is assumed to be known and is set to its true value. The simulations are carried out in GAUSS and the number of replications is 5000. Nominal significance level $\alpha = 0.05$ applies in all cases.

The performance of the CAIN test is compared to that of other p-value combination methods commonly applied in the literature. Using the p-values from the individual TSL test statistics as building blocks, we combine them into different panel statistics by employing (a) the standard inverse normal test without correction for the cross-sectional dependence; (b) both variants of Hartung's modified inverse normal test with κ_1 and κ_2, respectively, and (c) the multiple testing procedure of Simes (1986). For the latter test the individual p-values of the test statistics are ordered in ascending way as $p_{(1)} \leq \ldots \leq p_{(N)}$, and the joint null hypothesis $H_0 : r_i = r, \forall i$, is rejected at significance level α if $p_{(i)} \leq \frac{i\alpha}{N}$ for any $i = 1, \ldots, N$.

4.5.2 Simulation results

The size and power results under $H_0 : r_i = 0, \forall i$ when the true rank is zero and one, respectively, are presented in Table 4.3. In the following discussion average values of the parameters, computed over all replications, are denoted by a long bar.

With diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$, the cross-sectional dependence is very weak: $\overline{\hat{\rho}_\varepsilon} = 0.089$ and $\overline{\hat{\rho}_\varepsilon} = 0.068$ for $T = 100$ and $T = 200$, respectively, while the mean estimated correlation between the probits for the CAIN-TSL test is $\overline{\hat{\rho}_{\mathrm{t}}} = 0.001$. In this case all variants of the inverse normal test, including CAIN-TSL, become undersized as N grows. This results from the individual TSL tests being slightly undersized; the size distortions get magnified in the panel setting as N increases. Hartung's test with κ_1 has the most severe size distortion, inline with the findings of Demetrescu et al. (2006) that it is not suitable for more than $N = 5$ units when the cross-sectional correlation is low. The inverse normal and the CAIN-TSL tests perform best in this case both in terms of size and power.

Diagonal factor loading matrix with $U(0, 1)$ entries generates moderate cross-sectional correlation between the innovations: $\overline{\hat{\rho}_\varepsilon} = 0.18$, while the estimated average correlations

Table 4.3: Monte Carlo study results, not near-unit root processes (case A) and near-unit root processes (case B). Empirical size and power under $H_0 : r_i = 0$, $\forall i = 1, \ldots, N$.

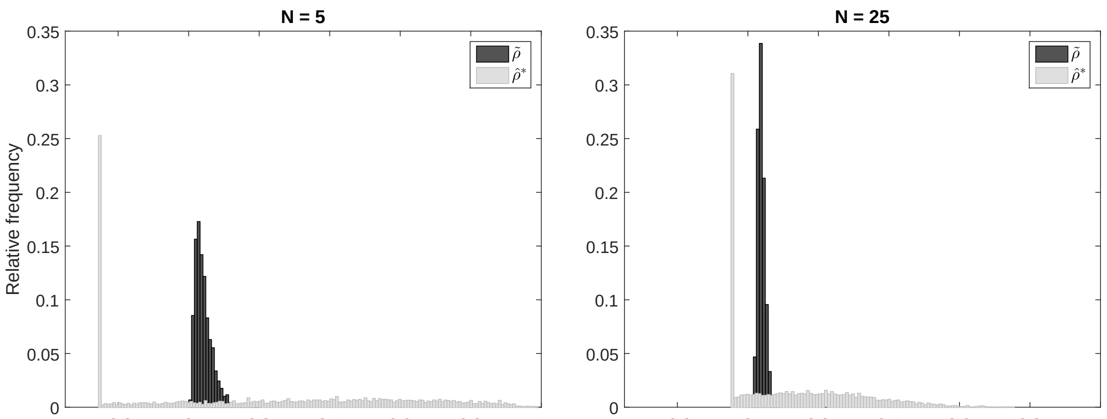
	True cointegrating rank 0: Size						True cointegrating rank 1, case A: Power						True cointegrating rank 1, case B: Power					
	T=100			T=200			T=100			T=200			T=100			T=200		
	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Inverse normal	0.05	0.04	0.03	0.04	0.04	0.03	0.56	0.92	0.99	0.99	1.00	1.00	0.16	0.27	0.36	0.58	0.92	0.99
Hartung κ_1	0.04	0.02	0.01	0.04	0.01	0.01	0.43	0.68	0.79	0.96	1.00	1.00	0.13	0.15	0.17	0.51	0.80	0.93
Hartung κ_2	0.05	0.04	0.03	0.05	0.03	0.02	0.46	0.73	0.82	0.97	1.00	1.00	0.16	0.23	0.28	0.54	0.85	0.95
Simes	0.06	0.06	0.07	0.05	0.05	0.06	0.36	0.50	0.57	0.94	1.00	1.00	0.14	0.17	0.19	0.48	0.68	0.78
CAIN-TSL	0.05	0.04	0.03	0.04	0.04	0.03	0.56	0.92	0.99	0.99	1.00	1.00	0.16	0.27	0.35	0.58	0.92	0.99
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.05	0.05	0.05	0.04	0.05	0.05	0.49	0.84	0.95	0.98	1.00	1.00	0.13	0.23	0.30	0.49	0.84	0.95
Hartung κ_1	0.04	0.02	0.01	0.04	0.02	0.01	0.35	0.56	0.66	0.93	1.00	1.00	0.10	0.12	0.12	0.42	0.65	0.79
Hartung κ_2	0.05	0.04	0.03	0.05	0.04	0.04	0.38	0.62	0.71	0.93	1.00	1.00	0.13	0.18	0.20	0.44	0.71	0.84
Simes	0.05	0.06	0.06	0.05	0.05	0.06	0.29	0.39	0.45	0.89	0.99	1.00	0.11	0.13	0.14	0.37	0.51	0.60
CAIN-TSL	0.04	0.04	0.04	0.04	0.04	0.04	0.48	0.83	0.93	0.98	1.00	1.00	0.13	0.21	0.26	0.49	0.82	0.94
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.06	0.08	0.11	0.06	0.09	0.12	0.42	0.71	0.82	0.95	1.00	1.00	0.13	0.22	0.29	0.42	0.69	0.82
Hartung κ_1	0.04	0.03	0.04	0.05	0.04	0.04	0.28	0.42	0.47	0.85	0.97	0.99	0.10	0.11	0.12	0.32	0.47	0.57
Hartung κ_2	0.06	0.06	0.06	0.06	0.06	0.07	0.31	0.47	0.53	0.86	0.97	0.99	0.12	0.16	0.17	0.35	0.54	0.64
Simes	0.06	0.06	0.06	0.05	0.05	0.06	0.24	0.31	0.33	0.79	0.93	0.96	0.10	0.11	0.11	0.28	0.37	0.42
CAIN-TSL	0.04	0.05	0.05	0.05	0.05	0.06	0.38	0.61	0.70	0.94	1.00	1.00	0.12	0.15	0.17	0.38	0.59	0.70
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.05	0.05	0.06	0.05	0.06	0.07	0.66	0.95	0.99	1.00	1.00	1.00	0.20	0.37	0.49	0.70	0.96	0.99
Hartung κ_1	0.04	0.02	0.02	0.04	0.02	0.02	0.52	0.77	0.87	0.98	1.00	1.00	0.17	0.23	0.27	0.62	0.90	0.97
Hartung κ_2	0.05	0.04	0.04	0.05	0.05	0.04	0.55	0.80	0.89	0.99	1.00	1.00	0.20	0.31	0.37	0.65	0.93	0.98
Simes	0.06	0.06	0.06	0.06	0.05	0.06	0.45	0.60	0.68	0.97	1.00	1.00	0.17	0.22	0.26	0.58	0.81	0.89
CAIN-TSL	0.04	0.03	0.03	0.04	0.04	0.04	0.63	0.92	0.98	1.00	1.00	1.00	0.18	0.29	0.35	0.67	0.93	0.99
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$						Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.07	0.12	0.15	0.08	0.14	0.18	0.68	0.94	0.98	0.99	1.00	1.00	0.26	0.44	0.54	0.73	0.95	0.98
Hartung κ_1	0.04	0.05	0.05	0.05	0.05	0.06	0.54	0.79	0.87	0.98	1.00	1.00	0.21	0.29	0.33	0.68	0.91	0.96
Hartung κ_2	0.06	0.07	0.08	0.06	0.07	0.08	0.57	0.82	0.89	0.98	1.00	1.00	0.23	0.36	0.41	0.70	0.93	0.98
Simes	0.05	0.05	0.06	0.05	0.05	0.05	0.49	0.67	0.73	0.97	1.00	1.00	0.20	0.28	0.32	0.65	0.87	0.93
CAIN-TSL	0.06	0.07	0.08	0.07	0.09	0.10	0.64	0.89	0.94	0.99	1.00	1.00	0.22	0.32	0.37	0.69	0.91	0.96

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice. Some results with size-adjusted power are available in Table 4.11 in the Appendix.

between the probits are already around $\overline{\tilde{\rho}_{\mathrm{t}}} = 0.007$ and $\overline{\tilde{\rho}_{\mathrm{t}}} = 0.005$ for $T = 100$ and $T = 200$, respectively. The inverse normal test has size close to the nominal one, along with the CAIN-TSL test, while both Hartung's tests again become undersized as N grows. Considering power, the inverse normal test performs best because the nominator of the panel test statistic is inflated by the unattended cross-sectional dependence. The CAIN-TSL test has slightly lower power compared to it for $T = 100$ in case B, but in case A and for $T = 200$ in case B it performs equally well.

Relatively strong cross-sectional correlation between the innovations is introduced by a diagonal factor loading matrix with $U(-1, 3)$ entries: $\overline{\rho_{\varepsilon}} = 0.413$, with estimated average correlation between the probits $\overline{\tilde{\rho}_{\mathrm{t}}}$ varying between 0.020 and 0.046. In this case the inverse normal test and Hartung's test with κ_2 become oversized for large N, while the size of Hartung's test with κ_1 and that of CAIN-TSL fluctuate around the desired 5% level. In terms of power, both in cases A and B the CAIN-TSL outperforms Hartung's tests and the test of Simes', being second only compared to the oversized standard inverse normal. The explanation for this power gain of CAIN-TSL in comparison to Hartung's tests is simple: it is due to the $\tilde{\rho}_{\mathrm{t}}$ estimator being much more precise than the $\hat{\rho}_{\mathrm{t}}^{*}$ one (see Figure 4.2). Therefore the variance of the test statistic does not get overestimated, which would lead to lower power. We note that due to the high correlation between the cross-sections, the power increase over N for all tests is less significant compared to the case of low cross-sectional correlation. This can be explained by the fact that the different cross-sections exhibit similar dynamics and in a sense carry very much the same information. Hence the marginal value (in terms of power) added by each individual unit is lower compared to when the cross-sectional dependence is weaker so that the information set gets richer as N grows.

Figure 4.2: Empirical distributions of Hartung's estimator $\hat{\rho}_{\mathrm{t}}^{*}$ and the CAIN estimator $\tilde{\rho}_{\mathrm{t}}$ for the correlation between the probits under $H_0 : r_i = 0, \forall i$, 3-variate VAR(2) process with a multifactor-error structure and a diagonal factor loading matrix
$\gamma_i \sim i.i.d.U(-1, 3)$, $T = 100$, 5000 replications



Lastly, with unrestricted factor loadings we investigate the performance of the tests when all variables are assumed to be correlated to the same extent, on average, over the cross-sections. When $\gamma_i \sim U(0,1)$, $\overline{\hat{\rho}_\varepsilon} = 0.36$ and the inverse normal test becomes oversized for high N, while both Hartung's κ_1 and the CAIN-TSL tests tend to become undersized. The CAIN-TSL test performs best in terms of power without becoming oversized for both cases A and B. When $\gamma_i \sim U(-1,3)$, $\overline{\hat{\rho}_\varepsilon} = 0.44$ the CAIN-TSL test tends to become oversized as N grows with size reaching 10% for $T = 200$ and $N = 25$ at the 5% level, while Hartung's κ_1 test observes the nominal size. As these experimental settings violate Assumption 2 for the CAIN-TSL test, we conclude that it is robust to such violations when the estimated mean absolute cross-sectional correlation between the innovations $\hat{\rho}_\varepsilon$ is low to moderate (≤ 0.35), or when the number of cross-sections is small. For higher estimated $\hat{\rho}_\varepsilon$ or when $N > 10$ we'd rather recommend to use Hartung's modification with $\kappa_1 = 0.2$.

Size and power under $H_0 : r_i = 1, \forall i$ for cases A and B are presented in Tables 4.7 and 4.9 in the Appendix, respectively. The undersized individual TSL test leads to all five panel tests being severely undersized. Their power nevertheless increases with N for T sufficiently large. For case A the inverse normal test exhibits the highest power, closely followed by the CAIN-TSL test. The same holds for case B when the cross-sectional correlation is low to moderate. For larger correlation the CAIN-TSL test has already slightly lower power than Hartung's tests. This is so because it precisely corrects for the high cross-sectional correlation between the probits, which turns out to be unnecessary given the severely undersized results of the individual TSL tests. In these cases the inverse normal test is again the most powerful test (without becoming oversized under the true null $H_0 : r_i = 1, \forall i$). In terms of size-adjusted power (see Table 4.11 in the Appendix) the CAIN test is exactly as powerful as the standard inverse normal.

These results can be summarized as follows. When there is only low mean absolute cross-sectional correlation $\hat{\rho}_\varepsilon$, in terms of size the CAIN-TSL performs comparably to the standard inverse normal and Hartung's κ_2 test, being as powerful as the inverse normal and more powerful than both Hartung's tests. This conclusion holds regardless of the null hypothesis under consideration. When the cross-sectional correlation is high, the CAIN-TSL test offers best size-power trade-off under the null hypothesis of no cointegration. This is particularly important in view of the sequential rank testing procedure, which begins with $H_0 : r_i = 0, \forall i$. Also, even when Assumption 2 is violated, the CAIN-TSL test may be used to test for no cointegration when the estimated $\hat{\rho}_\varepsilon$ is low to average, or when the number of cross-sections is not too large. Otherwise Hartung's test with κ_1 is recommended, as it controls the size better. When testing for higher cointegrating ranks under high cross-sectional correlation, then it might be preferable to employ the inverse normal test without any correction for the cross-sectional dependence, or Hartung's κ_2 test. They would yield high power without running the risk of an inflated Type I error probability, due to the lower than nominal size of the individual TSL test under such null hypotheses. Simes' simple intersection test, although almost correctly sized in all settings, is, in general, least powerful.

4.5.3 Comparison with Westerlund's (2006) test

To help position the CAIN-TSL test in the growing literature on panel cointegration, we evaluate its finite-sample properties in comparison with other similar tests. Our work is closest to that of Banerjee and Carrion-i Silvestre (2015), who propose a panel no-cointegration test accommodating level shifts, trend breaks and even breaks in the cointegrating relation. However, in the case of level and trend breaks their number and locations are restricted to be homogeneous over cross-sections, which may not always be the case in practice. Furthermore, their approach relies on first decomposing the observed time series into common and idiosyncratic components and determining their stochastic properties separately. It is, therefore, difficult to draw conclusions based on their test whether the observed variables actually exhibit cointegration or not, which is of primary interest in practice. Hence we rather compare the CAIN-TSL test with another test, which focuses on the cointegration properties of the observed variables.

Westerlund (2006) proposes an LM test for the null hypothesis of cointegration in panel data allowing for structural breaks in the deterministic terms of the data generating process. The heterogeneous number of breaks and the break dates are not necessarily known, but may be estimated from the data. The panel test statistic is computed as the normalized sum of the individual LM test statistics, where the first two moments of the statistic depend only on the number of breaks in each cross-section, but not on their locations or other nuisance parameters. Assuming independence of the cross-sectional units, Westerlund (2006) shows that the limiting distribution of the statistic is standard normal. For the general case of cross-sectionally dependent panels he proposes a bootstrapping algorithm. Nevertheless, he examines the performance of the test without bootstrap correction for a DGP featuring an unobserved common factor in the innovations, and he notes that it is "remarkably robust to moderate degrees of cross-sectional correlation". Hence it is this simpler version of the test which we employ next.

The comparison is based on the following DGP for Y_{it}, where the stochastic component $X_{it} = (X_{1t}, X_{2t})'$ follows the cointegrated VAR(2) process employed by Dolado and Lütkepohl (1996):

$$Y_{it} = \tilde{d}_{it} + X_{it}, \tag{4.19}$$

$$\Delta X_{it} = \begin{pmatrix} -\beta & \beta \\ 0 & 0 \end{pmatrix} X_{i,t-1} + \begin{pmatrix} 0.5 & 0.3 \\ 0 & 0.5 \end{pmatrix} \Delta X_{i,t-1} + u_{it}, \tag{4.20}$$

$$u_{it} = \gamma_i f_t + e_{it}. \tag{4.21}$$

The innovations u_{it} are augmented by a (2×1)-vector of unobserved common factors f_t such that $f_t \sim N(0, I_2)$. The factor loading matrix γ_i is diagonal with uniformly distributed entries $\sim U(-1, 3)$, while the idiosyncratic errors are generated as $e_{it} \sim N(0, I_2)$. The linear trend term $\tilde{d}_{it} = (d_{it}, 0)'$ affects only the first variable in the system, where d_{it} features one or two breaks with 50% chance each at individual-specific break fractions $\eta_{ij} \in (0.15T, 0.85T)$. In particular, $d_{it} = \delta_{ij} \cdot (1, t)'$, where the intercept and trend parameters are $\delta_{i1} \sim N(1, I_2)$ in the first sub-sample with $t < [\eta_{i1}T]$, $\delta_{i2} \sim N(2, I_2)$ in the second sub-sample with $t < [\eta_{i2}T]$, and $\delta_{i3} \sim N(0, I_2)$ in the third. To simulate cointegrated process with rank one, we set $\beta = 1$, while for no cointegration (rank is zero) $\beta = 0$. To put both tests on equal grounds, the lag order, the breaks number and the locations of the breaks are assumed to be known.

Table 4.4 summarizes the results. It is evident that although the CAIN-TSL test is undersized, its power increases significantly as N grows. Its size adjusted power is slightly lower than that of Westerlund's (2006) when N is as small as 5. However, Westerlund's (2006) test is severely oversized, despite the fact that the average absolute correlation $\bar{\hat{\rho}}_{\varepsilon}$ between the innovations for $i \neq j$ is only 0.3. Hence in practice the CAIN-TSL test offers a clear advantage in terms of better control of the Type-I error while maintaining good power properties.

Table 4.4: Comparison between the CAIN-TSL and Westerlund's (2006) test under factor-driven cross-sectional dependence

	Size			Power			Size-adjusted power		
$T \backslash N$	5	15	25	5	15	25	5	15	25
				CAIN-TSL					
100	0.007	0.000	0.000	0.636	0.867	0.938	0.764	0.967	0.994
200	0.020	0.007	0.009	0.813	0.979	0.998	0.858	0.990	0.999
				Westerlund's (2006) test					
100	0.491	0.974	0.999	0.997	1.000	1.000	0.937	1.000	1.000
200	0.393	0.935	0.993	1.000	1.000	1.000	1.000	1.000	1.000

Note: Rejection frequencies at the 5% nominal level based on 5000 replications. Power for Westerlund's (2006) test and size for the CAIN-TSL test are computed when the true rank and the hypothesized rank is zero. Size for Westerlund's (2006) test and power for the CAIN-TSL test are computed when the true cointegrating rank is zero, while the hypothesized rank is one.

4.6 Testing for cointegration between US regional house prices and personal income

We illustrate the use of the CAIN-TSL test by applying it to investigate a presumed cointegrating relationship between regional house prices and personal income in the United States. The evidence on the existence of a long-run relationship between house prices and macroeconomic fundamentals in the US housing market is rather mixed. On the one hand, using 23 years of quarterly data until 2002 on 95 metropolitan areas, Gallin (2006) fails to find support for a stable link between house prices and income. On the other hand, analyzing yearly data from 1975 to 2003 for 49 states, Holly et al. (2010) do establish a long-run equilibrium between real house prices and real per capita income with coefficients $(-1, 1)$, "as predicted by the theory".

We aim to shed some light on this issue by employing the CAIN-TSL test and the newest available data at quarterly frequency, spanning from 1983Q3 to 2018Q1. Such a long window with 141 time observations predisposes single-unit unit root and cointegration tests to have good finite-sample properties. The longer time series, however, come at a price. The global financial crisis of 2007-2008 has caused a structural break in many economic and financial time series, with personal income and house prices being no exception. Therefore, when trying to establish a cointegrating relationship between these two variables, this structural break has to be taken into account in order to avoid spurious inference.

In formulating the econometric model we follow Holly et al. (2010) and look at a log-linear relationship between real house prices, p_{it}, and the macroeconomic fundamentals real per capita disposable income, y_{it}, and real cost of borrowing net of house prices appreciation/depreciation, c_{it}. We model it in a VAR(p_i) framework as $Y_{it} = (p_{it}, y_{it}, c_{it})'$ for each state $i = 1, \ldots, N$, allowing for a level shift and trend break in the deterministic terms:

$$Y_{it} = \mu_{i0} + \mu_{i1}t + \delta_{i0}d_{it} + \delta_{i1}b_{it} + X_{it}, \quad t = 1, \ldots, T, \tag{4.22}$$

$$X_{it} = A_{1i}X_{i,t-1} + \ldots + A_{p_i,i}X_{i,t-p_i} + u_{it}. \tag{4.23}$$

The structural break is assumed to take place at a known date, and namely in the third quarter of 2007, which marks the beginning of the US mortgage and credit crisis. Therefore, $\tau = \eta T = 99$ with $\eta = 0.7$. Hence the dummy variables d_{it} and b_{it} take values $d_{it} = b_{it} = 0$ for $t < \tau$ and $d_{it} = 1$ and $b_{it} = t - \tau + 1$ for $t \geq \tau$. To account for potential lag effects, however, we check for robustness of this choice by shifting the break date by one or two quarters ahead.

The variables p_{it}, y_{it} and c_{it} are computed as in Holly et al. (2010). We use state-level data for all 50 states and the District of Columbia (N=51). Information on the data sources is listed in Table 4.12 in the Appendix. Although available, data for eight years prior to 1983 has not been included into the analysis in order to avoid issues arising from heteroscedasticity due to the Great Moderation. The time span of the remaining data is nevertheless sufficiently large.

Each variable is first tested for a unit roots by a univariate ADF test. Perron (1989) has shown that not accounting for an existing structural break reduces the test's ability to reject a false null hypothesis of a unit root, so any non-rejections would require further investigation. Rejections, however, would point to stationarity. To reflect their trending behaviour we allow for an intercept and a linear time trend in the models for the income and house prices variables, while for net cost of borrowing only an intercept is included. The results are presented in the first panel of Table 4.13 in the Appendix. For y_{it} and p_{it} we observe rejections of the unit root null at 5% significance level for only two and four states, respectively. In contrast, c_{it} is largely classified as stationary with rejections at the 5% level for 37 states and the largest p-value not exceeding 0.14. These results are in line with those of Holly et al. (2010). We therefore next focus only on the house prices and personal income variables.

To investigate whether the income and house price variables are nonstationary also when a structural break is allowed for in the level and the slope of the DGP, we apply the test of Popp (2008). His innovational-outlier (IO) type test is suitable for the application at hand as it allows the break to gradually take effect and permits a break to occur under both the null and the alternative hypotheses. Furthermore, the limiting distribution of the test statistic with an endogenous selection of the break date is the same as in the case of a known break, as is the case here. We again choose to employ a single-unit unit root test, as deciding upon the stochastic properties of the variables on a unit-by-unit basis and not at the panel level allows us to exclude stationary time series from the subsequent cointegration analysis. Tables 4.14 and 4.15 in the Appendix summarize the results. Allowing for a structural break around the time of the beginning of the financial crisis does indeed result in rejection of the unit root null at 5% in some cases.

Excluding these few states where either variable is deemed stationary, we proceed to testing for cointegration with the CAIN-TSL test. The lag order of each individual VAR process has been selected by the modified AIC (MAIC) criterion of Qu and Perron (2007), whose computation has been augmented by dummy variables to account for the structural break. Table 4.5 presents the results. We note that for a break located in 2007Q3 not a single rejection of the no-cointegration null is observed. If the break is set to one quarter later, a single rejection at the 5% significance level emerges, and namely for North Carolina. For a break in 2008Q1 the null of zero cointegrating rank is rejected at 5% only for North Dakota. Turning to the CAIN-TSL test, we first note that the estimated average absolute correlation between the innovations to the same variables in the panel, $\hat{\rho}_\varepsilon$, is about 0.42. It is comparable to the magnitude of the correlations of the variables in first differences between regions reported by Holly et al. (2010) and points to the existence of significant cross-sectional dependence, as expected. This value of $\hat{\rho}_\varepsilon$ leads to an estimated correlation between the probits $\tilde{\rho}_{\mathrm{t}}$ of about 0.05. The estimate of the average absolute pairwise correlation between the innovations to the different variables is 0.039 and below, pointing that Assumption 2 is not violated. The highly positive CAIN-TSL statistics, far from the rejection region in the left tail of the standard normal distribution, only further corroborate the lack of cointegration between real house prices and real per capita disposable income in the selected period.

These findings are in stark contrast with those of Holly et al. (2010), who do establish that cointegration exists at the panel level. An important difference to their study is our extended sample. Holly et al. (2010) use data until 2003 and do not cover the subsequent housing bubble in the early 2000s, which affected more than twenty US states and began to collapse in the mid-2000s, leading to the subprime mortgage crisis[5]. Our results suggest that the house price dynamics have diverged from the evolution of the real per capita disposable income for most, if not all US states in this period.

One question, however, poses further interest. Is there evidence in the data for another, newer departure of the house prices from the income fundamental in the period following the burst of the housing bubble? To answer it, we apply the intersection-type panel cointegration test by Arsova and Örsal (2019) to the data sample from 2008Q1 to 2018Q1 ($T = 41$). The p_{it} and y_{it} variables are found to be $I(1)$ for most states by the ADF test (see the right panel of Table 4.13 in the Appendix). The real net cost of borrowing c_{it} is, as in the full sample, mostly $I(0)$; the results are not reported to save space. No major economic event has led to a structural break in this period, hence we employ the single-unit cointegration test of Saikkonen and Lutkepohl (2000) without breaks for each state with nonstationary house prices and income variables. The individual p-values are then combined in a panel test by the Simes' procedure. The results are presented in Table 4.16 in the Appendix. Allowing for a trend in the variables but not in the error-correction (EC) term leads to a rejection of the no-cointegration null hypothesis in 13 states at the 5% level. Two of these values are smaller than the corresponding critical values of Simes, leading to a rejection at the 5% level of the composite null hypothesis of no-cointegration for all states. Hence we find evidence that the US real house prices have not (yet) departed from the equilibrium relationship with real disposable income in the period following the burst of the last house prices bubble.

[5]Hu and Oxley (2018) find evidence that following the dot-com crisis many regional house price bubbles have formed on a state level within different time frames instead of a single synchronous divergence from the intrinsic value of the house prices.

Table 4.5: Results of the CAIN-TSL test with structural breaks for US house prices dataset

	Break in 2007Q3			Break in 2007Q4			Break in 2008Q1		
	$\hat{p}_i$	$LR^{\text{trace}}_{\text{TSL}}$	$p-$value	$\hat{p}_i$	$LR^{\text{trace}}_{\text{TSL}}$	$p-$value	$\hat{p}_i$	$LR^{\text{trace}}_{\text{TSL}}$	$p-$value
Alabama	5	5.81	0.930	4	5.61	0.940	–	–	–
Arizona	7	8.57	0.719	4	7.75	0.794	4	8.11	0.761
Arkansas	–	–	–	–	–	–	–	–	–
California	7	13.02	0.301	7	13.54	0.262	7	15.47	0.151
Colorado	7	9.40	0.637	7	8.50	0.725	7	8.04	0.767
Connecticut	3	8.94	0.682	2	6.74	0.874	2	6.31	0.902
Delaware	5	6.19	0.910	5	6.55	0.887	–	–	–
District of Columbia	–	–	–	–	–	–	–	–	–
Florida	–	–	–	–	–	–	–	–	–
Georgia	5	3.93	0.989	8	15.36	0.157	7	17.60	0.077*
Idaho	–	–	–	8	11.67	0.412	–	–	–
Illinois	–	–	–	–	–	–	–	–	–
Indiana	8	5.40	0.949	8	6.65	0.880	7	9.34	0.640
Iowa	8	9.73	0.603	8	8.26	0.747	7	7.40	0.822
Kansas	6	6.45	0.894	6	7.56	0.811	6	11.16	0.458
Kentucky	8	9.57	0.619	8	10.39	0.535	–	–	–
Louisiana	8	4.26	0.983	8	4.71	0.973	8	6.51	0.889
Maine	–	–	–	–	–	–	–	–	–
Maryland	5	8.21	0.753	4	7.59	0.808	4	7.97	0.773
Massachusetts	4	10.06	0.569	5	9.00	0.675	4	9.17	0.657
Michigan	8	5.36	0.951	8	6.25	0.906	7	6.38	0.897
Minnesota	8	5.17	0.958	8	6.91	0.862	7	8.60	0.714
Mississippi	8	8.01	0.772	8	11.12	0.464	8	12.27	0.358
Missouri	7	6.45	0.894	7	7.64	0.804	7	11.17	0.457
Montana	6	7.39	0.825	6	8.26	0.747	7	3.97	0.988
Nebraska	7	9.68	0.608	8	11.88	0.393	7	12.74	0.320
Nevada	–	–	–	–	–	–	–	–	–
New Hampshire	6	9.75	0.600	5	11.73	0.406	4	8.78	0.696
New Jersey	2	6.15	0.912	4	5.17	0.958	6	5.90	0.925
New Mexico	5	8.06	0.767	5	11.66	0.413	4	15.65	0.143
New York	5	6.16	0.912	4	5.28	0.954	4	5.04	0.962
North Carolina	8	9.04	0.672	8	20.55	0.027**	–	–	–
North Dakota	7	9.92	0.583	5	15.59	0.147	7	19.29	0.043**
Ohio	8	6.43	0.895	7	8.89	0.687	7	11.04	0.470
Oklahoma	5	7.29	0.833	6	3.96	0.988	5	7.49	0.816
Oregon	–	–	–	–	–	–	–	–	–
Pennsylvania	5	14.23	0.219	6	18.49	0.057*	5	16.76	0.101
Rhode Island	5	10.05	0.570	4	11.16	0.460	4	10.47	0.526
South Carolina	6	8.55	0.721	8	2.99	0.997	–	–	–
South Dakota	8	13.40	0.273	8	14.47	0.203	8	13.75	0.246
Tennessee	5	5.10	0.961	4	4.52	0.978	8	9.15	0.659
Texas	7	6.86	0.866	7	7.27	0.834	6	9.95	0.578
Utah	–	–	–	–	–	–	–	–	–
Vermont	5	9.83	0.593	4	9.29	0.647	4	9.63	0.611
Virginia	8	8.35	0.740	8	11.30	0.446	8	11.91	0.389
Washington	8	13.52	0.265	8	12.35	0.353	–	–	–
West Virginia	–	–	–	–	–	–	–	–	–
Wisconsin	8	7.70	0.799	8	8.41	0.733	7	8.63	0.711
Wyoming	7	7.63	0.806	6	6.53	0.888	6	11.83	0.396
Alaska	8	15.65	0.145	8	12.63	0.330	8	15.82	0.136
Hawaii	8	6.75	0.874	8	7.93	0.777	6	8.91	0.684

		CAIN-TSL	
$t(\tilde{\rho})$	2.603	1.818	0.723
$\hat{\rho}_\varepsilon$	0.426	0.421	0.416
$\hat{\rho}_\varepsilon{}^o$	0.039	0.038	0.038
$\tilde{\rho}_t$	0.055	0.054	0.052

Notes: $LR^{\text{trace}}_{\text{TSL}}$ denotes the LR trace statistic of Trenkler et al. (2007). The lag order $\hat{p}_i$ is selected by the MAIC criterion of Qu and Perron (2007), augmented to account for the structural break. $\hat{\rho}_\varepsilon$ denotes the average absolute pairwise cross-sectional correlation between the innovations to the same variables in the panel. $\hat{\rho}_\varepsilon{}^o$ denotes the average absolute pairwise cross-sectional correlation between the innovations to the different variables in the panel. $\tilde{\rho}_t$ denotes the estimated correlation between the individual probits.

4.7 Conclusion

In this paper we propose a new meta-analytic approach (CAIN) to test for the cointegrating rank in panels where structural breaks and cross-sectional dependence are allowed for. It is an extension of the likelihood-based rank test of Trenkler et al. (2007) (TSL), and requires only the p-values of the individual LR trace statistics of the TSL test. The CAIN-TSL test is based on a modification of the popular inverse normal method for p-values combination, employing a novel estimator for the unknown correlation between the probits. We propose a way to estimate this correlation as a function of the system dimension, the cointegrating rank under the null hypothesis and the average absolute cross-sectional correlation between the residuals of the individual VAR models in first differences. The latter is easily estimable in practice and provides an easy-to-interpret measure of the degree of cross-sectional dependence. In a Monte Carlo study we demonstrate the superior properties of the CAIN-TSL test in comparison with other meta-analytic approaches recently proposed in the panel unit-root literature. An application of the test to investigate a presumed long-run equilibrium relationship between house prices and personal income in a panel of 51 US states provides an illustration of its usefulness in practice.

Bibliography

Arsova, A., Örsal, D.D.K., 2019. Intersection tests for the cointegrating rank in dependent panel data. Forthcoming in Communications in Statistics: Simulation and Computation , doi:10.1080/03610918.2018.1489552.

Arsova, A., Örsal, D.D.K., 2018. Likelihood-Based Panel Cointegration Test in the Presence of a Linear Time Trend and Cross-Sectional Dependence. Econometric Reviews 37, 1033–1050.

Banerjee, A., Carrion-i Silvestre, J.L., 2015. Cointegration in Panel Data with Structural Breaks and Cross-Section Dependence. Journal of Applied Econometrics 30, 1–23.

Choi, I., 2001. Unit root tests for panel data. Journal of International Money and Finance 20, 249–272.

Costantini, M., Lupi, C., 2013. A Simple Panel-CADF Test for Unit Roots. Oxford Bulletin of Economics and Statistics 75, 276–296.

Demetrescu, M., Hanck, C., 2012. Unit Root Testing in Heteroscedastic Panels Using the Cauchy Estimator. Journal of Business & Economic Statistics 30, 256–264.

Demetrescu, M., Hassler, U., Tarcolea, A.I., 2006. Combining Significance of Correlated Statistics with Application to Panel Data. Oxford Bulletin of Economics and Statistics 68, 647–663.

Dolado, J.J., Lütkepohl, H., 1996. Making wald tests work for cointegrated VAR systems. Econometric Reviews 15, 369–386.

Gallin, J., 2006. The long run relationship between house prices and income: evidence from local housing markets. Real Estate Economics 34, 417–438.

Hanck, C., 2013. An intersection test for panel unit roots. Econometric Reviews 32, 183–203.

Hartung, J., 1999. A Note on Combining Dependent Tests of Significance. Biometrical Journal 41, 849–855.

Holly, S., Pesaran, H.M., Yamagata, T., 2010. A spatio-temporal model of house prices in the USA. Journal of Econometrics 158, 160–173.

Hu, Y., Oxley, L., 2018. Bubbles in US regional house prices: evidence from house price–income ratios at the State level. Applied Economics doi:10.1080/00036846.2017.1418080.

Johansen, S., Mosconi, R., Nielsen, B., 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. Econometrics Journal 3, 216–249.

Maddala, G.S., Wu, S., 1999. A comparative study of panel unit root tests. Oxford Bulletin of Economics and Statistics Special Issue, 631–652.

Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361–1401.

Pesaran, M.H., 2015. Testing Weak Cross-Sectional Dependence in Large Panels. Econometric Reviews 34, 1089–1117.

Popp, S., 2008. New innovational outlier unit root test with a break at an unknown time. Journal of Statistical Computation and Simulation 78, 1145–1161.

Qu, Z., Perron, P., 2007. A modified information criterion for cointegration tests based on a VAR approximation. Econometric Theory 23, 638–685.

Saikkonen, P., Lutkepohl, H., 2000. Trend Adjustment Prior to Testing for the Cointegrating Rank of a Vector Autoregressive Process. Journal of Time Series Analysis 21, 435–456.

Sarkar, S.K., 1998. Some Probability Inequalities for Ordered MTP2 Random Variables: A Proof of the Simes Conjecture. The Annals of Statistics 26, 494–504.

Simes, R.J., 1986. An improved Bonferroni procedure for multiple tests of significance. Biometrika 73, 751–754.

Stouffer, S., DeVinney, L., Suchmen, E., 1949. The American Soldier: Adjustment During Army Life. Princeton University Press, Princeton, NJ.

Toda, H.Y., 1995. Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions. Econometric Theory 11, 1015–1032.

Trenkler, C., Saikkonen, P., Lütkepohl, H., 2007. Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break. Journal of Time Series Analysis 29, 331–358.

Wagner, M., Hlouskova, J., 2010. The Performance of Panel Cointegration Methods: Results from a Large Scale Simulation Study. Econometric Reviews 29, 182–223.

Westerlund, J., 2006. Testing for Panel Cointegration with Multiple Structural Breaks. Oxford Bulletin of Economics and Statistics 68, 101–132.

Westerlund, J., Edgerton, D.L., 2008. A Simple Test for Cointegration in Dependent Panels with Structural Breaks. Oxford Bulletin of Economics and Statistics 70, 665–704.

4.A Appendix

Figure 4.3: Cross-sectional correlation of the probits against cross-sectional correlation of the innovations, bi-variate system with true cointegrating rank zero, testing for $H_0 : r_i = 0, \forall i$

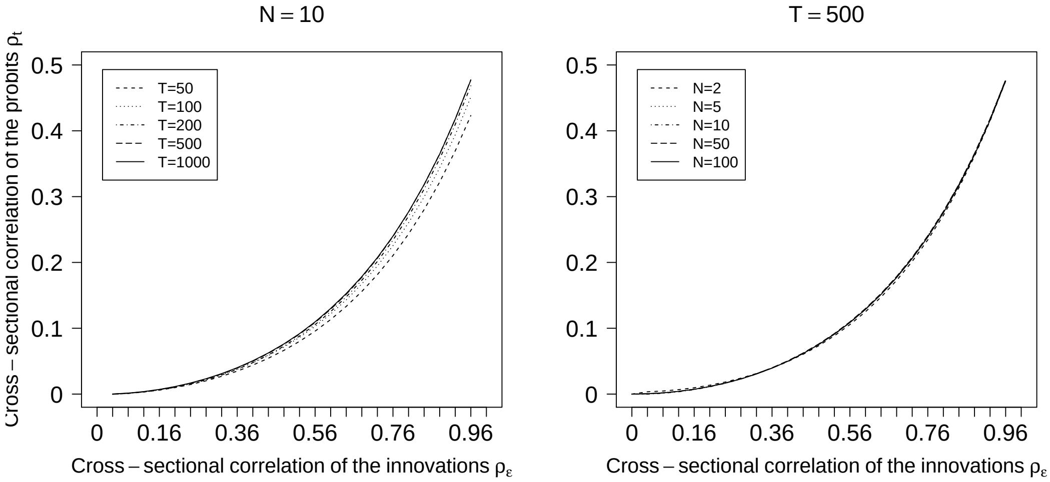


Table 4.6: True and estimated ρ_ε under $H_0 : r = 0$

	Estimated $\hat{\rho}_\varepsilon$		
True ρ_ε	True rank 0	True rank 1	True rank 2
0.2	0.200	0.196	0.192
0.4	0.400	0.392	0.384
0.6	0.600	0.589	0.577
0.8	0.800	0.786	0.771

Notes: Results from large-scale simulation study with $m = 3$, $T = 500$ and $N = 5$.

Table 4.7: Monte Carlo study results, case A: not near-unit root processes. Empirical size under $H_0 : r_i = 1$, $\forall i$, $i = 1, \ldots, N$.

	T=100			T=200		
	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Inverse normal	0.001	0.000	0.000	0.003	0.000	0.000
Hartung κ_1	0.001	0.000	0.000	0.005	0.000	0.000
Hartung κ_2	0.002	0.000	0.000	0.007	0.001	0.000
Simes	0.010	0.006	0.009	0.016	0.017	0.016
CAIN-TSL	0.001	0.000	0.000	0.003	0.000	0.000
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.002	0.000	0.000	0.006	0.001	0.000
Hartung κ_1	0.004	0.000	0.000	0.006	0.001	0.000
Hartung κ_2	0.005	0.001	0.000	0.008	0.001	0.000
Simes	0.009	0.007	0.004	0.020	0.019	0.024
CAIN-TSL	0.002	0.000	0.000	0.005	0.001	0.000
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.004	0.003	0.002	0.015	0.017	0.017
Hartung κ_1	0.003	0.001	0.001	0.014	0.007	0.005
Hartung κ_2	0.004	0.002	0.001	0.018	0.012	0.011
Simes	0.008	0.007	0.005	0.024	0.018	0.019
CAIN-TSL	0.002	0.001	0.000	0.010	0.008	0.005
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.003	0.002	0.002	0.009	0.007	0.006
Hartung κ_1	0.003	0.001	0.001	0.008	0.004	0.003
Hartung κ_2	0.005	0.002	0.002	0.011	0.007	0.005
Simes	0.011	0.010	0.007	0.019	0.014	0.018
CAIN-TSL	0.002	0.001	0.000	0.006	0.002	0.002
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.003	0.002	0.002	0.009	0.007	0.006
Hartung κ_1	0.003	0.001	0.001	0.008	0.004	0.003
Hartung κ_2	0.005	0.002	0.002	0.011	0.007	0.005
Simes	0.011	0.010	0.007	0.019	0.014	0.018
CAIN-TSL	0.002	0.001	0.000	0.006	0.002	0.002

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice. Some results with size-adjusted power are available in Table 4.11 in the Appendix.

Table 4.8: Monte Carlo study results, case A: not near-unit root processes. Empirical power under $H_0 : r_i = 1$, $\forall i$, $i = 1, \ldots, N$.

	T=100			T=200		
	$N=5$	$N=15$	$N=25$	$N=5$	$N=15$	$N=25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Inverse normal 0.41	0.78	0.92	0.99	1.00	1.00	
Hartung κ_1	0.28	0.46	0.56	0.97	1.00	1.00
Hartung κ_2	0.32	0.54	0.63	0.98	1.00	1.00
Simes	0.23	0.27	0.27	0.95	1.00	1.00
CAIN-TSL	0.41	0.77	0.91	0.99	1.00	1.00
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.40	0.75	0.88	0.99	1.00	1.00
Hartung κ_1	0.26	0.42	0.51	0.97	1.00	1.00
Hartung κ_2	0.28	0.49	0.58	0.97	1.00	1.00
Simes	0.20	0.24	0.25	0.95	1.00	1.00
CAIN-TSL	0.39	0.72	0.85	0.99	1.00	1.00
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.40	0.67	0.79	0.99	1.00	1.00
Hartung κ_1	0.24	0.34	0.41	0.95	1.00	1.00
Hartung κ_2	0.27	0.40	0.47	0.96	1.00	1.00
Simes	0.18	0.21	0.21	0.92	0.99	1.00
CAIN-TSL	0.34	0.55	0.63	0.98	1.00	1.00
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.43	0.78	0.90	0.99	1.00	1.00
Hartung κ_1	0.30	0.49	0.60	0.98	1.00	1.00
Hartung κ_2	0.33	0.56	0.67	0.98	1.00	1.00
Simes	0.24	0.29	0.31	0.96	1.00	1.00
CAIN-TSL	0.40	0.69	0.81	0.99	1.00	1.00
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.45	0.73	0.82	0.99	1.00	1.00
Hartung κ_1	0.30	0.46	0.53	0.97	1.00	1.00
Hartung κ_2	0.33	0.52	0.58	0.97	1.00	1.00
Simes	0.24	0.29	0.32	0.94	0.99	1.00
CAIN-TSL	0.39	0.60	0.68	0.99	1.00	1.00

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice. Some results with size-adjusted power are available in Table 4.11 in the Appendix.

Table 4.9: Monte Carlo study results, case B: near-unit root processes. Empirical size under $H_0 : r_i = 1$, $\forall i$, $i = 1, \ldots, N$.

	T=100			T=200		
	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Inverse normal	0.000	0.000	0.000	0.001	0.000	0.000
Hartung κ_1	0.000	0.000	0.000	0.002	0.000	0.000
Hartung κ_2	0.000	0.000	0.000	0.003	0.000	0.000
Simes	0.005	0.003	0.002	0.008	0.005	0.006
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.000	0.000	0.000	0.002	0.000	0.000
Hartung κ_1	0.001	0.000	0.000	0.001	0.000	0.000
Hartung κ_2	0.001	0.000	0.000	0.002	0.000	0.000
Simes	0.004	0.002	0.001	0.008	0.005	0.006
CAIN-TSL	0.000	0.000	0.000	0.002	0.000	0.000
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.000	0.000	0.000	0.003	0.001	0.000
Hartung κ_1	0.000	0.000	0.000	0.002	0.000	0.000
Hartung κ_2	0.001	0.000	0.000	0.003	0.001	0.001
Simes	0.003	0.002	0.001	0.008	0.004	0.005
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.000	0.000	0.000	0.002	0.000	0.000
Hartung κ_1	0.000	0.000	0.000	0.001	0.000	0.000
Hartung κ_2	0.001	0.000	0.000	0.002	0.000	0.000
Simes	0.004	0.003	0.002	0.008	0.006	0.006
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.000	0.000	0.000	0.002	0.001	0.001
Hartung κ_1	0.000	0.000	0.000	0.003	0.001	0.001
Hartung κ_2	0.001	0.000	0.000	0.004	0.001	0.001
Simes	0.005	0.003	0.004	0.009	0.005	0.006
CAIN-TSL	0.000	0.000	0.000	0.001	0.000	0.000

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice. Some results with size-adjusted power are available in Table 4.11 in the Appendix.

Table 4.10: Monte Carlo study results, case B: near-unit root processes. Empirical power under $H_0 : r_i = 1$, $\forall i$, $i = 1, \ldots, N$.

	T=100			T=200		
	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$
Test	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-0.4, 0.4)$					
Inverse normal	0.02	0.01	0.01	0.37	0.66	0.83
Hartung κ_1	0.03	0.01	0.00	0.34	0.55	0.70
Hartung κ_2	0.03	0.02	0.01	0.38	0.65	0.81
Simes	0.04	0.03	0.04	0.33	0.46	0.53
CAIN-TSL	0.02	0.01	0.01	0.37	0.66	0.83
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.03	0.02	0.01	0.38	0.67	0.82
Hartung κ_1	0.02	0.01	0.00	0.32	0.53	0.66
Hartung κ_2	0.03	0.02	0.01	0.36	0.63	0.76
Simes	0.04	0.03	0.03	0.30	0.41	0.47
CAIN-TSL	0.03	0.01	0.01	0.37	0.64	0.79
	Diagonal factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.04	0.05	0.05	0.39	0.64	0.76
Hartung κ_1	0.02	0.02	0.02	0.31	0.47	0.58
Hartung κ_2	0.03	0.03	0.04	0.34	0.55	0.65
Simes	0.03	0.03	0.03	0.27	0.35	0.40
CAIN-TSL	0.03	0.02	0.02	0.34	0.52	0.61
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(0, 1)$					
Inverse normal	0.03	0.02	0.02	0.39	0.66	0.80
Hartung κ_1	0.03	0.01	0.01	0.37	0.57	0.71
Hartung κ_2	0.04	0.02	0.01	0.41	0.67	0.81
Simes	0.05	0.05	0.04	0.37	0.52	0.62
CAIN-TSL	0.02	0.01	0.01	0.35	0.57	0.68
	Unrestricted factor loading matrix $\gamma_i \sim i.i.d.U(-1, 3)$					
Inverse normal	0.05	0.04	0.06	0.41	0.64	0.75
Hartung κ_1	0.04	0.02	0.02	0.38	0.56	0.67
Hartung κ_2	0.05	0.04	0.04	0.42	0.64	0.76
Simes	0.05	0.05	0.05	0.39	0.54	0.61
CAIN-TSL	0.04	0.02	0.02	0.37	0.51	0.60

Notes: Rejection frequencies at 5% significance level, 5000 replications. Power is not size-adjusted, as size-adjustment would not be available in practice. Some results with size-adjusted power are available in Table 4.11 in the Appendix.

Table 4.11: Size-adjusted power for true rank one under $H_0 : r_i = 0, \forall i$

	$T = 100$			$T = 200$		
	$N = 5$	$N = 15$	$N = 25$	$N = 5$	$N = 15$	$N = 25$
Diagonal factor loadings $\gamma_i \sim i.i.d.U(-1, 3)$						
Inverse normal	0.40	0.62	0.70	0.95	1.00	1.00
Hartung κ_1	0.30	0.49	0.54	0.85	0.98	0.99
Hartung κ_2	0.29	0.43	0.47	0.82	0.96	0.98
CAIN-TSL	0.40	0.62	0.70	0.94	1.00	1.00
Unrestricted factor loadings $\gamma_i \sim i.i.d.U(-1, 3)$						
Inverse normal	0.62	0.84	0.90	0.97	1.00	1.00
Hartung κ_1	0.57	0.81	0.86	0.97	1.00	1.00
Hartung κ_2	0.54	0.77	0.80	0.97	1.00	1.00
CAIN-TSL	0.62	0.85	0.90	0.99	1.00	1.00
Unrestricted factor loadings $\gamma_i \sim i.i.d.U(-1, 3)$, near-unit root processes						
Inverse normal	0.20	0.25	0.28	0.67	0.85	0.91
Hartung κ_1	0.22	0.31	0.32	0.67	0.90	0.95
Hartung κ_2	0.21	0.31	0.32	0.66	0.90	0.95
CAIN-TSL	0.20	0.25	0.28	0.66	0.85	0.91

Rejection frequencies at 5% nominal level, 5000 replications.

Table 4.12: US house prices dataset: data description

Variable	Description	ID	Source
$P_{it,g}$	CPI_{CA} for California; CPI_{FL} for Florida; CPI_{GA} for Georgia; CPI_{TX} for Texas; CPI_{MI} for Michigan; CPI_{IL} for Illinois, Indiana and Wisconsin; CPI_{MA} for Massatchusetts and New Hampshire; CPI_{PA} for Delaware and Maryland; CPI_{NY} for New York, New Jersey and Pennsylvania; CPI_{West} for Alaska, Arizona, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oklahoma, Oregon, Utah, Washington and Wyoming; CPI_{US} for all remaining states.		
CPI_{US}	Consumer Price Index: Total All Items for the United States; Index 2010=100, Seasonally Adjusted	CPALTT01USQ661S	OECD
CPI_{CA}	Consumer Price Index for All Urban Consumers: All items in San Francisco-Oakland-Hayward, CA (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA422SA0	U.S. Bureau of Labor Statistics
CPI_{FL}	Consumer Price Index for All Urban Consumers: All items in Miami-Fort Lauderdale-West Palm Beach, FL (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA320SA0	U.S. Bureau of Labor Statistics
CPI_{GA}	Consumer Price Index for All Urban Consumers: All items in Atlanta-Sandy Springs-Roswell, GA (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA319SA0	U.S. Bureau of Labor Statistics
CPI_{TX}	Consumer Price Index for All Urban Consumers: All items in Houston-The Woodlands-Sugar Land, TX (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA318SA0	U.S. Bureau of Labor Statistics
CPI_{MI}	Consumer Price Index for All Urban Consumers: All items in Detroit-Warren-Dearborn, MI (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA208SA0	U.S. Bureau of Labor Statistics
CPI_{IL}	Consumer Price Index for All Urban Consumers: All items in Chicago-Naperville-Elgin, IL-IN-WI (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA207SA0	U.S. Bureau of Labor Statistics
CPI_{MA}	Consumer Price Index for All Urban Consumers: All items in Boston-Cambridge-Newton, MA-NH (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA103SA0	U.S. Bureau of Labor Statistics
CPI_{PA}	Consumer Price Index for All Urban Consumers: All items in Philadelphia-Camden-Wilmington, PA-NJ-DE-MD (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA102SA0	U.S. Bureau of Labor Statistics
CPI_{NY}	Consumer Price Index for All Urban Consumers: All items in New York-Newark-Jersey City, NY-NJ-PA (CBSA); Index 1982-1984=100; Not Seasonally Adjusted	CUURA101SA0	U.S. Bureau of Labor Statistics
CPI_{West}	Consumer Price Index for All Urban Consumers: All items in West; Index 1982-1984=100; Not SA	CUUR0400SA0	U.S. Bureau of Labor Statistics
$P_{it,h}$	All-Transactions House Price Index; Index 1980:Q1=100; Not Seasonally Adjusted	**STHPI	U.S. Federal Housing Finance Agency
PD_{it}	Approximated by the difference $PInc_{it}$ - G_{it}		
$PInc_{it}$	United States, BEA, Personal Income by Major Component (SQ4), Personal Income, USD		Macrobond
G_{it}	United States, BEA, Personal Income by Major Component (SQ4), Employee and Self-Employed Contributions for Government Social Insurance, USD		Macrobond
POP_{it}	Population in number of persons by state; United States, BEA, Personal Income Summary: Personal Income, Population, per Capita Personal Income (CA1), Population, States; Population (Midperiod) is used from 2017Q1 to 2018Q1; linear interpolation to yield quarterly data, carried out by Macrobond		Macrobond
RB_t	Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for the United States	IRLTLT01USQ156N	OECD
p_{it}	Natural logarithm of the US State real house price index, $p_{it} = \ln(P_{it,h}/P_{it,g})$		
y_{it}	Natural logarithm of the US State real per capita disposable income, $y_{it} = \ln(PD_{it}/(POP_{it} \times P_{it,g}))$		
r_{it}	US State real long-term interest rate, $r_{it} = RB_t/100 - ln(P_{it,g}/P_{I,t-1,g})$		
c_{it}	US State real cost of borrowing net of real house price appreciation/depreciation, $c_{it} = r_{it} - \Delta p_{it}$		

Notes: OECD stands for OECD, Main Economic Indicators - complete database. Where available, ID denotes the variable ID in FRED (Federal Reserve Economic Data). ** stands for the two-letter abbreviation of the corresponding state. Not seasonally adjusted CPI series have been seasonally adjusted by the Census X12 method. All indices have been rebased to 1980Q1=100. Quarterly CPI indices have been derived from monthly and bi-monthly series by taking average values of the three months in a quarter.

Table 4.13: ADF unit root test results for US house prices dataset

	Sample 1983Q1 – 2018Q1 (T=141)									Sample 2008Q1 – 2018Q1 (T=41)					
	y_{it}			p_{it}			c_{it}			y_{it}			p_{it}		
	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value
Alabama	3	−2.690	0.243	8	−1.232	0.899	3	−4.676	0.0002***	1	−2.860	0.186	3	−0.746	0.962
Arizona	4	−1.702	0.746	7	−1.494	0.827	3	−3.112	0.028**	4	−2.910	0.170	3	−1.341	0.863
Arkansas	4	−4.106	0.008***	4	−1.031	0.936	3	−3.276	0.018**	4	−2.744	0.225	4	−3.764	0.029**
California	4	−2.121	0.530	4	−0.823	0.960	4	−2.571	0.101	6	−2.994	0.146	5	−5.748	0.0001***
Colorado	4	−1.267	0.892	7	−1.742	0.727	8	−3.134	0.026**	4	−1.947	0.612	4	−1.280	0.879
Connecticut	1	−1.666	0.761	4	−1.694	0.749	3	−3.379	0.013**	1	−2.900	0.173	3	−2.190	0.482
Delaware	5	−1.489	0.829	8	−1.203	0.906	3	−3.552	0.008***	3	−2.394	0.377	1	−0.884	0.948
District of Columbia	3	−1.935	0.631	5	−0.764	0.966	4	−4.662	0.0002***	6	−1.941	0.615	1	−3.731	0.031**
Florida	4	−2.092	0.545	5	−0.831	0.960	3	−2.936	0.044**	5	−2.725	0.232	3	−4.132	0.012**
Georgia	3	−2.516	0.320	7	−0.884	0.954	3	−3.528	0.009***	3	−2.370	0.389	4	−1.002	0.933
Idaho	4	−1.626	0.778	7	−1.635	0.774	4	−3.970	0.002***	4	−2.733	0.230	3	−1.280	0.879
Illinois	3	−2.865	0.177	8	−0.928	0.949	3	−2.931	0.044**	4	−2.830	0.195	3	−1.680	0.742
Indiana	1	−3.414	0.054*	7	−1.026	0.936	4	−2.712	0.074*	6	−2.613	0.277	3	−1.007	0.932
Iowa	5	−2.887	0.170	7	−1.482	0.831	4	−2.645	0.087*	5	−1.846	0.664	3	−1.290	0.876
Kansas	4	−2.146	0.515	5	−1.625	0.778	4	−2.851	0.054*	5	−1.709	0.729	3	−0.829	0.954
Kentucky	3	−2.277	0.443	8	−0.993	0.941	4	−2.817	0.059*	5	−2.771	0.216	3	−0.669	0.969
Louisiana	4	−1.729	0.733	5	−3.667	0.028**	7	−2.895	0.048**	3	−2.753	0.222	1	−1.421	0.840
Maine	3	−2.168	0.503	6	−1.844	0.678	3	−4.106	0.001***	3	−2.239	0.457	3	−1.281	0.879
Maryland	3	−1.764	0.717	6	−0.701	0.971	3	−2.894	0.049**	3	−1.478	0.821	3	−2.799	0.206
Massachusetts	4	−2.588	0.287	7	−2.374	0.391	3	−3.033	0.034**	3	−2.729	0.231	3	−2.228	0.462
Michigan	4	−3.301	0.070*	8	−0.939	0.948	3	−3.670	0.006***	5	−2.928	0.165	4	−2.562	0.299
Minnesota	5	−2.468	0.344	8	−0.883	0.954	8	−3.457	0.011**	4	−2.741	0.226	4	−1.871	0.651
Mississippi	2	−1.398	0.858	6	−1.605	0.787	5	−3.804	0.004***	6	−4.305	0.008***	3	−1.055	0.924
Missouri	6	−2.597	0.282	8	−0.960	0.945	3	−2.505	0.117	4	−3.335	0.075*	3	−0.923	0.943
Montana	1	−2.918	0.160	8	−1.474	0.834	7	−6.917	0.000***	1	−1.658	0.752	3	−1.450	0.830

Notes: Intercept and trend included for p_{it} and y_{it}, only intercept included for c_{it}. $\hat{p}_i$ denotes the lag order chosen by AIC with a maximum of 8 lags for the 1983Q1–2018Q1 sample and with a maximum of 6 lags for the 2008Q1–2018Q1 sample, respectively. We are grateful to Christoph Hanck for providing us with the GAUSS code for the computation of the $p-$values.

ADF unit root test results for US house prices dataset (continued)

	Sample 1983Q1 – 2018Q1 (T=141)									Sample 2008Q1 – 2018Q1 (T=41)					
	y_{it}			p_{it}			c_{it}			y_{it}			p_{it}		
	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value	$\hat{p}_i$	t-stat	$p-$value
Nebraska	1	−3.222	0.084*	8	−1.474	0.834	6	−2.404	0.142	1	−1.783	0.695	3	−1.038	0.927
Nevada	5	−1.289	0.887	6	−0.811	0.961	3	−3.750	0.004***	1	−3.097	0.121	4	−3.998	0.017**
New Hampshire	5	−2.190	0.491	7	−1.370	0.866	3	−3.073	0.031**	6	−3.047	0.133	3	−1.833	0.670
New Jersey	3	−2.137	0.521	5	−1.574	0.798	3	−3.039	0.034**	1	−3.011	0.142	3	−2.257	0.447
New Mexico	4	−1.588	0.793	6	−1.971	0.612	3	−3.511	0.009***	4	−2.378	0.385	3	−1.410	0.843
New York	2	−3.249	0.079*	6	−2.151	0.513	3	−3.651	0.006***	1	−4.028	0.015*	3	−1.539	0.799
North Carolina	1	−2.815	0.194	8	−1.158	0.914	3	−2.766	0.066*	1	−2.450	0.350	3	−0.396	0.984
North Dakota	5	−2.244	0.461	7	−2.098	0.542	6	−8.977	0.000***	1	−0.404	0.984	5	−1.291	0.876
Ohio	3	−3.419	0.053*	8	−0.894	0.953	3	−3.159	0.025**	6	−2.749	0.223	4	−1.153	0.907
Oklahoma	5	−1.810	0.695	4	−4.464	0.002***	3	−2.804	0.060*	4	−1.011	0.931	1	−1.421	0.840
Oregon	3	−2.131	0.524	5	−1.182	0.910	5	−3.722	0.005***	4	−2.281	0.434	6	−2.119	0.520
Pennsylvania	2	−3.295	0.071*	6	−1.304	0.883	3	−2.577	0.100	5	−3.005	0.143	3	−1.312	0.871
Rhode Island	4	−2.213	0.478	5	−1.569	0.801	5	−3.659	0.006***	1	−3.066	0.128	3	−1.511	0.809
South Carolina	1	−3.021	0.130	7	−1.206	0.905	3	−3.649	0.006***	4	−2.533	0.312	3	−0.482	0.981
South Dakota	1	−2.186	0.494	8	−3.389	0.057*	7	−11.940	0.000***	6	−2.542	0.308	3	−0.927	0.943
Tennessee	1	−2.698	0.239	8	−1.188	0.909	4	−4.151	0.001***	6	−2.630	0.270	3	−0.595	0.974
Texas	4	−1.498	0.826	4	−1.161	0.914	3	−2.479	0.123	1	−1.776	0.698	4	−0.920	0.944
Utah	4	−1.885	0.657	5	−1.650	0.768	5	−3.351	0.015**	4	−2.614	0.276	4	−1.909	0.632
Vermont	3	−2.081	0.552	6	−0.859	0.957	3	−4.142	0.001***	6	−2.992	0.147	5	−1.487	0.818
Virginia	4	−1.841	0.680	5	−0.610	0.977	3	−2.921	0.046**	4	−2.686	0.247	3	−2.406	0.371
Washington	2	−3.069	0.118	4	−0.935	0.948	4	−2.521	0.113	4	−2.673	0.253	4	−1.472	0.823
West Virginia	1	−2.947	0.151	7	−3.306	0.070*	4	−7.782	0.000***	3	−3.348	0.073*	1	−1.815	0.679
Wisconsin	3	−3.177	0.093*	8	−0.603	0.977	4	−2.434	0.134	6	−3.293	0.082*	3	−0.745	0.963
Wyoming	3	−2.147	0.515	6	−4.539	0.002***	5	−5.230	0.000***	3	−1.684	0.740	1	−1.410	0.843
Alaska	1	−3.667	0.028**	8	−2.392	0.382	7	−7.267	0.000***	4	−1.690	0.738	1	−2.175	0.490
Hawaii	4	−2.317	0.422	7	−3.705	0.025**	3	−2.705	0.076*	3	−1.845	0.664	1	−1.838	0.668

Notes: Intercept and trend included for p_{it} and y_{it}, only intercept included for c_{it}. $\hat{p}_i$ denotes the lag order chosen by AIC with a maximum of 8 lags for the 1983Q1–2018Q1 sample and with a maximum of 6 lags for the 2008Q1–2018Q1 sample, respectively. We are grateful to Christoph Hanck for providing us with the GAUSS code for the computation of the $p-$values.

Table 4.14: Results of Popp's (2008) unit root test with structural breaks for US Real per capita disposable income y_{it}

Break in:	2007Q3		2007Q4		2008Q1	
	$\hat{p}_i$	t-stat	$\hat{p}_i$	t-stat	$\hat{p}_i$	t-stat
Alabama	0	−4.11*	7	−3.43	2	−4.65**
Arizona	3	−3.22	3	−3.21	3	−4.01*
Arkansas	3	−6.38	3	−5.96**	3	−7.18**
California	3	−3.19	3	−3.19	3	−3.58
Colorado	3	−4.06*	3	−3.66	3	−3.78
Connecticut	4	−3.43	4	−3.34	4	−3.17
Delaware	8	−2.59	8	−2.27	8	−2.20
District of Columbia	3	−4.65**	3	−4.57**	3	−4.90**
Florida	2	−4.33**	3	−4.39**	3	−4.80**
Georgia	2	−3.21	2	−2.39	2	−2.43
Idaho	8	−4.30**	8	−4.18*	8	−4.85**
Illinois	2	−4.31**	2	−4.12*	2	−4.78**
Indiana	0	−2.60	2	−2.86	2	−3.25
Iowa	4	−2.80	4	−2.83	4	−2.86
Kansas	3	−3.25	3	−3.23	0	−1.84
Kentucky	2	−3.22	2	−3.23	3	−4.28**
Louisiana	7	−2.79	7	−3.16	3	−3.95*
Maine	2	−2.32	2	−2.29	2	−2.40
Maryland	2	−2.01	2	−2.12	2	−2.16
Massachusetts	3	−2.96	3	−3.03	3	−3.08
Michigan	8	−2.31	8	−1.99	3	−3.02
Minnesota	4	−2.67	4	−2.90	4	−3.14
Mississippi	1	−3.34	1	−3.29	1	−3.96*
Missouri	3	−3.72	3	−3.71	3	−4.19*
Montana	0	−2.26	0	−2.39	0	−2.30
Nebraska	0	−3.38	0	−3.39	0	−3.44
Nevada	4	−5.63**	4	−4.71**	4	−4.30**
New Hampshire	2	−2.15	2	−2.16	3	−2.55
New Jersey	2	−2.63	2	−2.85	2	−2.95
New Mexico	7	−2.16	7	−2.27	7	−2.36
New York	0	−3.26	0	−3.37	0	−3.38
North Carolina	0	−3.17	8	−3.37	2	−4.42**
North Dakota	4	−2.01	4	−1.56	4	−1.31
Ohio	2	−3.13	2	−3.21	2	−3.79
Oklahoma	4	−1.66	4	−1.94	4	−1.57
Oregon	2	−3.56	2	−3.59	2	−4.33**
Pennsylvania	0	−3.37	0	−3.44	2	−3.89
Rhode Island	2	−2.58	2	−2.64	2	−2.73
South Carolina	3	−3.84	0	−3.67	3	−4.82**
South Dakota	6	−3.27	6	−2.88	6	−2.37
Tennessee	0	−2.65	0	−2.62	0	−2.9
Texas	2	−3.04	2	−3.18	3	−3.91
Utah	3	−4.16*	3	−4.28**	3	−5.60**
Vermont	2	−2.60	2	−2.68	2	−2.75
Virginia	3	−2.74	3	−2.87	3	−3.09
Washington	5	−4.11*	5	−4.16	5	−4.62**
West Virginia	0	−5.12**	0	−5.06**	0	−4.80**
Wisconsin	0	−3.36	0	−3.24	2	−3.16
Wyoming	2	−3.16	2	−3.36	2	−2.84
Alaska	8	−3.46	8	−3.41	8	−2.52
Hawaii	2	−1.88	2	−2.00	2	−2.21

Notes: A single break in both intercept and trend included. Optimal lag order denoted by $\hat{p}_i$. The 5% and 10% critical values are −4.2736 and −3.9417, respectively, for a break in 2007Q3; −4.2638 and −3.9312, respectively, for a break in 2007Q4; and −4.2536 and −3.9203, respectively, for a break in 2008Q1.

Table 4.15: Results of Popp's (2008) unit root test with structural breaks for US Real house price index p_{it}

Break in:	2007Q3		2007Q4		2008Q1	
	$\hat{p}_i$	t-stat	$\hat{p}_i$	t-stat	$\hat{p}_i$	t-stat
Alabama	6	−1.68	6	−1.77	6	−1.96
Arizona	0	−1.79	0	−1.93	0	−2.05
Arkansas	3	−3.52	3	−3.53	3	−3.39
California	3	−3.70	3	−3.67	3	−3.32
Colorado	6	−3.47	6	−3.48	6	−3.33
Connecticut	3	−3.17	3	−3.24	3	−3.13
Delaware	7	−3.90	7	−4.13*	7	−4.32**
District of Columbia	4	−2.59	4	−2.79	4	−2.89
Florida	3	−3.13	3	−3.11	6	−3.74
Georgia	6	−3.30	6	−3.34	6	−3.34
Idaho	6	−3.42	5	−3.20	6	−3.69
Illinois	7	−4.50**	7	−4.62**	7	−4.64**
Indiana	6	−1.56	8	−1.80	8	−1.52
Iowa	8	−2.99	8	−2.92	8	−2.64
Kansas	4	−1.94	4	−2.04	6	−2.16
Kentucky	0	−1.35	0	−1.40	7	−2.18
Louisiana	0	−2.81	0	−3.03	0	−3.16
Maine	8	−4.49**	8	−4.38**	8	−4.39**
Maryland	4	−2.66	4	−2.87	6	−3.97*
Massachusetts	5	−3.90	3	−3.24	5	−3.78
Michigan	6	−3.26	6	−3.29	6	−2.55
Minnesota	7	−3.81	7	−3.83	7	−3.62
Mississippi	0	−1.74	0	−1.83	0	−1.94
Missouri	7	−3.36	7	−3.46	7	−3.56
Montana	7	−3.34	7	−3.54	7	−3.59
Nebraska	6	−2.18	6	−2.15	6	−1.86
Nevada	3	−4.05*	3	−3.80	3	−3.44
New Hampshire	6	−3.79	6	−3.86	6	−3.80
New Jersey	3	−3.32	3	−3.34	3	−3.24
New Mexico	3	−1.96	3	−2.36	3	−2.38
New York	5	−3.80	5	−3.74	5	−3.65
North Carolina	6	−3.20	6	−3.29	6	−3.44
North Dakota	6	−2.22	6	−2.44	6	−2.36
Ohio	3	−0.58	3	−0.64	6	−1.07
Oklahoma	3	−3.64	3	−3.72	3	−3.72
Oregon	6	−4.66**	4	−4.57**	4	−4.44**
Pennsylvania	5	−3.52	5	−3.54	5	−3.46
Rhode Island	4	−3.77	4	−3.78	3	−3.20
South Carolina	6	−1.72	6	−1.80	6	−2.03
South Dakota	7	−2.49	7	−2.56	7	−2.54
Tennessee	6	−2.98	6	−3.06	6	−3.13
Texas	3	−2.12	3	−2.29	3	−2.33
Utah	6	−4.58**	4	−4.31**	6	−4.48**
Vermont	5	−3.58	5	−3.73	5	−3.55
Virginia	3	−2.43	3	−2.51	6	−3.93
Washington	3	−3.11	3	−3.35	3	−3.29
West Virginia	4	−1.72	4	−2.00	4	−2.09
Wisconsin	2	−1.87	2	−1.99	2	−1.70
Wyoming	3	−3.12	3	−3.32	3	−3.50
Alaska	7	−2.68	7	−2.83	7	−2.95
Hawaii	3	−3.51	3	−3.56	3	−3.57

Notes: A single break in both intercept and trend included. Optimal lag order denoted by $\hat{p}_i$. The 5% and 10% critical values are −4.2736 and −3.9417, respectively, for a break in 2007Q3; −4.2638 and −3.9312, respectively, for a break in 2007Q4; and −4.2536 and −3.9203, respectively, for a break in 2008Q1.

Table 4.16: Results of the Simes-SL test for US house prices dataset, sample 2008Q1–2018Q1

	Trend allowed in the EC term			Trend orthogonal to the EC term			
	$\hat{p}_i$	$LR_{\text{SL}}^{\text{trace}}$	$p-$value	$\hat{p}_i$	$LR_{\text{SL}}^{\text{trace}}$	$p-$value	Simes' crit. value
Utah	7	14.39	0.083	7	25.85	0.000*	0.001
Washington	4	18.28	0.018	4	18.41	0.001*	0.002
Louisiana	8	6.63	0.704	8	14.42	0.006	0.003
North Dakota	8	19.62	0.010	8	14.65	0.006	0.005
Idaho	2	15.24	0.060	2	14.25	0.007	0.006
Oklahoma	3	10.94	0.259	3	13.82	0.009	0.007
Oregon	4	12.47	0.161	4	12.92	0.013	0.008
Wyoming	3	13.72	0.105	3	12.68	0.015	0.009
South Carolina	2	13.74	0.104	2	10.84	0.033	0.010
Rhode Island	2	8.26	0.517	2	10.79	0.034	0.011
Texas	4	10.58	0.288	4	10.34	0.041	0.013
Colorado	7	11.89	0.194	7	10.19	0.044	0.014
Missouri	2	8.81	0.456	2	10.00	0.048	0.015
Montana	2	11.74	0.203	2	9.50	0.059	0.016
North Carolina	6	8.32	0.510	6	8.66	0.083	0.017
Arizona	2	15.90	0.047	1	7.78	0.119	0.018
Kentucky	2	8.41	0.500	2	7.53	0.132	0.019
Illinois	2	10.89	0.263	2	7.31	0.143	0.020
Minnesota	2	9.18	0.417	2	7.26	0.146	0.022
New Hampshire	2	6.06	0.767	2	7.06	0.158	0.023
Tennessee	2	9.59	0.376	2	6.81	0.174	0.024
Virginia	2	6.00	0.774	1	6.79	0.175	0.025
Ohio	2	7.14	0.646	2	6.69	0.182	0.026
New Jersey	2	7.11	0.650	2	6.65	0.185	0.027
West Virginia	1	10.66	0.281	1	6.58	0.190	0.028
New Mexico	2	6.42	0.728	2	6.53	0.194	0.030
Alaska	1	6.27	0.745	1	6.39	0.204	0.031
Kansas	1	5.92	0.782	1	6.34	0.207	0.032
Massachusetts	2	6.27	0.745	2	6.16	0.222	0.033
Wisconsin	2	7.77	0.574	2	5.91	0.243	0.034
Pennsylvania	2	7.18	0.642	2	5.75	0.258	0.035
Connecticut	2	5.83	0.792	2	5.61	0.272	0.036
Vermont	2	6.71	0.696	1	5.43	0.290	0.038
Georgia	2	6.86	0.679	2	5.39	0.294	0.039
Alabama	2	7.98	0.548	2	5.25	0.308	0.040
Hawaii	5	12.96	0.137	2	5.13	0.322	0.041
Michigan	4	10.40	0.303	4	5.03	0.333	0.042
South Dakota	2	5.39	0.835	2	4.62	0.381	0.043
Nebraska	2	7.32	0.625	2	4.58	0.387	0.044
Delaware	2	6.11	0.762	2	4.31	0.422	0.045
Iowa	2	7.58	0.595	7	4.04	0.461	0.047
Maine	2	3.99	0.940	2	3.75	0.504	0.048
Indiana	2	5.66	0.809	2	3.01	0.624	0.049
Maryland	2	4.74	0.891	2	2.81	0.658	0.050

Notes: $LR_{\text{SL}}^{\text{trace}}$ denotes the LR trace statistic of Saikkonen and Lutkepohl (2000). The lag order $\hat{p}_i$ is selected by the MAIC criterion of Qu and Perron (2007). The results are presented sorted by the $p-$value of the test with trend orthogonal to the EC term, as required for the Simes procedure. p-values which are smaller than the corresponding critical value of Simes are marked by an asterisk. Arkansas, California, District of Columbia, Florida, Mississippi, Nevada and New York are excluded as the unit root null hypothesis has been rejected by the ADF test at the 5% level for either y_{it} or p_{it} or both.

5

Exchange rate pass-through to import prices in Europe: A panel cointegration approach

Antonia Arsova

This paper takes a panel cointegration approach to the estimation of short- and long-run exchange rate pass-through (ERPT) to import prices in the European countries. Although economic theory suggests a long-run relationship between import prices and exchange rate, in recent empirical studies its existence has either been overlooked, or it has proven difficult to establish. Resorting to novel tests for panel cointegration, we find support of the equilibrium relationship hypothesis. Exchange rate pass-through elasticities, estimated by two different techniques for cointegrated panel regressions, give insight into the most recent development of the ERPT.

Keywords: Exchange rate pass-through, import prices, panel cointegration, cross-sectional dependence, common factors
JEL classification: C12, C23, F31

5.1 Introduction

Exchange rate pass-through measures the extent to which import prices, expressed in the currency of the importing country, reflect changes in the exchange rate with its trading partners. Assuming that export prices are determined by a markup over marginal costs, the import price elasticity w.r.t. the exchange rate depends on the exporters' pricing strategies. If exporters choose to absorb exchange rate fluctuations into their markup, a strategy also known as local currency pricing (LCP) or pricing-to-market, then import prices remain largely unaffected by exchange rate shocks and the ERPT is said to be incomplete. On the other hand, if exporters choose not to adjust their markup, exchange rate fluctuations get reflected in full into import prices, which is known as producer currency pricing (PCP). The ERPT in this case is said to be complete. Under complete ERPT, depreciation of the importing country's currency translates into increase of import prices and may thus lead to inflation. Therefore, the degree and the determinants of exchange rate pass-through into import prices (a.k.a. first-stage pass-through) and subsequently into consumer prices (second-stage pass-through) pose an important issue to policy-makers looking to stabilize inflation, especially in a monetary union such as the euro area.

In the ever growing body of empirical literature on ERPT one issue becomes apparent – namely, whether there exists a long-run equilibrium relationship between import prices, nominal exchange rate and other potential macroeconomic determinants of import prices. For example, Campa and Goldberg (2005), Ben Cheikh and Rault (2016) and Ben Cheikh and Rault (2017) find no or only weak evidence of cointegration and proceed to estimate an ERPT equation in first differences. De Bandt and Razafindrabe (2014) do not even consider the possibility of cointegrating relations and having established the nonstationarity of the model variables proceed to estimate a model in first differences as well.

On the other hand, De Bandt et al. (2008) and Brun-Aguerre et al. (2012) do establish a cointegrating relation and thus estimate error-correction (EC) models for the ERPT. However, De Bandt et al. (2008) allow for level shifts and structural breaks in the cointegrating relation, while Brun-Aguerre et al. (2012) employ individual-unit and first-generation panel cointegration tests, whose results might be compromised by unattended cross-sectional dependence. Delatte and López-Villavicencio (2012) and Brun-Aguerre et al. (2017) also find strong evidence for cointegration, but they focus on asymmetric ERPT – that is, allowing the effects of exchange rate appreciation or depreciation on import prices to differ. Consequently, they argue that imposing the restriction of symmetric ERPT may hinder revealing the long-run equilibrium. The evidence on the existence of a linear cointegration relationship is, therefore, inconclusive.

The presence or absence of cointegration determines the choice of estimation methodology and models which do not consider it have been criticized on two grounds. First, ignoring a significant error-correction (EC) term leads to omitting essential information and hence to inferior model performance (Brun-Aguerre et al., 2012). Second, by evading the notion of cointegration as long-run equilibrium, other ad-hoc measures of long-run ERPT need to be constructed, whose estimates strongly depend on the choice of other model parameters, e.g. the lag order, and can thus become unreliable (De Bandt et al., 2008). Therefore, the debate on whether cointegration underlies the ERPT not only constitutes an interesting econometric puzzle but has far-reaching consequences concerning the estimation results.

The contribution of this paper is twofold. First, employing novel second-generation panel cointegration tests it provides evidence on the existence of a long-run equilibrium relationship between import prices and nominal exchange rate for a panel of nineteen European countries. Contrary to some recent findings (e.g. De Bandt et al., 2008), cointegration emerges without the necessity to allow for structural breaks neither in the deterministic terms, nor in the cointegrating relation. The cointegrating relationship is shown to be driven by unobserved global stochastic trends. Second, taking cointegration and its driving forces into account, the paper presents estimates of the long-run and short-run pass-through elasticities at the panel level and for the individual countries using most recent data covering the period since the introduction of the Euro in 1999. The continuously updated fully modified (Cup-FM), and continuously updated bias-corrected (Cup-BC) estimators of Bai et al. (2009), and the dynamic common correlated effects (DCCE) estimator of Chudik and Pesaran (2015), all of which are robust to cross-sectional dependence induced by unobserved common factors, are employed. Despite the technical differences of these estimators, the results they yield are remarkably similar. Following a 1% depreciation of the exchange rate, the import prices are inclined to rise by 0.37% on average as estimated by the Cup-FM and Cup-BC estimators, and by 0.33% as estimated by the DCCE estimator. These results indicate only partial pass-through, rejecting both the LCP and PCP hypotheses for the panel as a whole.

The rest of the paper is organized as follows. Section 2 postulates the econometric model for the ERPT and describes the data used for the analysis. Section 3 presents the results of the unit root and cointegration analyses. Section 4 describes the econometric methodology for the estimation and discusses the empirical results, and Section 5 concludes. Auxiliary results are collected in the Appendix.

5.2 Model and data

5.2.1 Exchange rate pass-through into import prices

The analysis is based on the framework adopted by Campa and Goldberg (2005), which is commonly applied in the literature. For notational simplicity the model is written suppressing the dependence on the cross-sectional dimension i. It assumes that the import prices, P_t, equal the export prices of the country's trading partners, P_t^x, multiplied by the exchange rate, E_t, expressed per unit of foreign currency:

$$P_t = E_t P_t^x. \tag{5.1}$$

The export prices comprise the producers' marginal cost, C_t, and gross markup, M_t:

$$P_t^x = C_t M_t. \tag{5.2}$$

The marginal cost, in turn, depends on the wages in the exporting market, W_t, and on the demand conditions in the importing market, Y_t. Denoting the logarithms of all variables by lowercase letters, eq. (5.1) thus becomes

$$\begin{aligned} p_t &= e_t + c_t + m_t \\ &= e_t + a_1 y_t + a_2 w_t + m_t. \end{aligned} \tag{5.3}$$

The markup is assumed to comprise both a fixed effect ϕ and a component depending the macroeconomic conditions, which may be reflected in the exchange rate and/or the demand conditions:

$$m_t = \phi + b_1 e_t + b_2 y_t. \tag{5.4}$$

Hence the general ERPT equation in log-linear form becomes

$$p_t = \phi + (1 + b_1)e_t + (a_1 + b_2)y_t + a_2 w_t, \tag{5.5}$$

or, more succinctly,

$$p_t = \beta_0 + \beta_1 e_t + \beta_2 y_t + \beta_3 w_t. \tag{5.6}$$

The primary focus of this paper is the pass-through elasticity given by the coefficient β_1 in eq. (5.6). If $\beta_1 = 1$, the pass-through to import prices is said to be complete. Exchange rate fluctuations are reflected one-to-one in the exporters' prices in the domestic market, and in this case producer currency pricing is present. If $\beta_1 = 0$, then exchange rate movements do not affect the prices in the importing market. Exporters do not adjust their prices abroad, but rather fully absorb the exchange rate fluctuations in their markup, and hence local currency pricing takes place.

5.2.2 Data description

The dataset comprises a balanced panel ($T = 77, N = 19$) with quarterly time series covering the period 1999Q1 – 2018Q1 for nineteen European countries: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Italy, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

The data on import prices are taken from the Main Economic Indicators (MEI) database of the OECD and reflect the prices of non-commodity imports of goods and services. Nominal effective exchange rate (NEER), weighted by the unit labour costs of a country's trading partners, is taken from the IMF International Financial Statistics (IFS) database for the model's exchange rate variable. It is defined in quantity notation such that an increase represents an appreciation of the domestic currency. This implies that the coefficient β_1 in (5.6) is expected to be negative, with $\beta_1 = -1$ indicating complete pass-through. Domestic demand is approximated by real GDP taken from the OECD Quarterly National Accounts database.

The choice of variable for the producers' costs is more involved, since there exists no directly observed variable which controls for the trade shares of the exporting countries. Therefore, a proxy for w has to be constructed from trade data. We follow Bailliu and Fujii (2004), who exploit the real effective exchange rate (REER) based on unit labour costs to create a trade-weighted measure of foreign producers' costs. Denoting the natural logarithm of REER by q, it can be represented as

$$q_t = e_t + ulc_t - ulc_t^*, \tag{5.7}$$

where ulc_t and ulc_t^* stand for the domestic and foreign unit labour costs in natural logarithms, respectively. REER is given in price notation, such that an increase reflects a worsening of the international competitive position, and e is given in quantity notation. Solving eq. (5.7) for ulc_t^* yields a trade weighted proxy for foreign producers' costs, which is then taken as w in the analysis. The unit labour costs series are obtained from the OECD MEI database, while REER and NEER are taken from IMF IFS.

5.3 Preliminary analysis

5.3.1 Testing for cross-sectional dependence

The first step of the analysis is to determine the degree and source of cross-sectional dependence in the panel. This is important in order to select the correct tools for analysing the integration and cointegration properties of the data and for the subsequent estimation of the ERPT. It is well-known that unattended strong cross-sectional dependence may result in oversized panel unit root and cointegration tests and biased estimates of the slope coefficients in eq. (5.6) (see, e.g., Banerjee et al., 2004 and Phillips and Sul, 2003, 2007).

For this aim the CD test of Pesaran (2015) is applied to the panel with country cross-sections for each variable in eq. (5.6). The test assumes weak[1] cross-sectional dependence under the null hypothesis, such as a spatial-type dependence or dependence driven by common factors affecting only a limited number of units as $N \to \infty$, for example. Rejection of the null is taken as evidence of the presence of strong cross-sectional dependence such as one caused by global (unobserved) common factors. The test statistic is computed as the standardized average of the pairwise correlation coefficients between the series in the panel and is normally distributed under the null hypothesis. To avoid spurious correlation arising from unit roots, the variables have been transformed into first differences. The results are presented in Table 5.1.

Table 5.1: Pesaran's (2015) *CD* statistic for the observed data

Variable	CD test statistic	p-value	$\overline{\hat{\rho}_{ij}}$	$\left\lvert\overline{\hat{\rho}_{ij}}\right\rvert$
Δp	42.72***	0.000	0.375	0.390
Δe	42.95***	0.000	0.377	0.483
Δy	52.45***	0.000	0.460	0.461
Δw	12.03***	0.000	0.106	0.172

Notes: $\overline{\hat{\rho}_{ij}}$ denotes the average pairwise correlation coefficient while $\left\lvert\overline{\hat{\rho}_{ij}}\right\rvert$ denotes the average absolute pairwise correlation coefficient over cross-sections.
*, ** and *** denote significance at the 10, 5 and 1% level, respectively.

The null of weak cross-sectional dependence is convincingly rejected for all variables. This is expected, given the tight economic and financial links between the European countries and the common currency and monetary policy in the euro area. Hence the analysis proceeds taking into account the presence of strong cross-sectional dependence.

5.3.2 Unit root and cointegration analysis

Unit root testing

Next the integration and cointegration properties of the time series are examined by second-generation panel unit root tests which are robust to cross-sectional dependence. In particular, the simple panel unit root test of Pesaran (2007) and the meta-analytic tests of Demetrescu et al. (2006) and Hanck (2013) are applied to the panel with

[1]For definitions of notions of weak and strong cross-sectional dependence refer to Chudik et al. (2011).

country cross-sections of each variable in eq. (5.6)[2]. Tables 5.8 and 5.9 in the Appendix summarize the results. The test of Pesaran (2007) rejects only for the exchange rate series at lags 1, 2, 3 and 4 and for the import price series at lag 1 (Table 5.9). On the other hand, a unit root at the panel level cannot be rejected for any variable in levels by the tests of Hanck (2013) and Demetrescu et al. (2006) (Table 5.8). All three tests reject the presence of a unit root in the first-differenced variables.[3] Hence there is prevailing evidence of the presence of unit roots in all variables in the model.

Cointegration testing

The next step in the analysis is to test the observed variables for cointegration. For this purpose the meta-analytic test of Arsova and Örsal (2019) is employed. Similarly to the panel unit root test of Hanck (2013), this test is too based on Simes' multiple testing procedure, where p-values from individual-unit likelihood-ratio (LR) cointegrating rank tests of Saikkonen and Lütkepohl (2000) (SL) are used. Two versions of the latter test are considered, one allowing for a deterministic time trend both in the variables in levels and in the error-correction (EC) term, and one allowing for a trend only in the variables in levels only. Denoting the cointegrating rank of the system for country i by r_i, the null hypothesis of the test is $H_0 : r_i = r$, where $r = 0, 1, 2, 3$ denotes the common cointegrating rank in a sequential testing procedure. The alternative hypothesis is $H_1 : r_i > r$ for at least one i.

The results are presented in Table 5.2. As the smallest individual p-value for testing $H_0 : r = 0$ by the first variant of the SL test is lower than the corresponding Simes' critical value, while $H_0 : r = 1$ cannot be rejected, there is evidence of a single cointegrating relationship in the panel at the 5% significance level. In order to ensure that the long-run equilibrium connects not only a certain pair of variables, the test of Arsova and Örsal (2019) is applied to all eight different bi-variate systems. The null hypothesis of no cointegration is rejected for neither pair; the results are omitted for brevity. Hence the equilibrium relationship is more complex, involving at least three or all four of the variables in the system.

Analysis of the unobserved common and idiosyncratic components

Having established nonstationarity and the presence of a single long-run equilibrium relationship in the data, the analysis proceeds to uncover their driving forces. For this purpose the approach of panel analysis of nonstationarity in idiosyncratic and

[2] More details on the computation of the tests by Hanck (2013) and Demetrescu et al. (2006) are given in the Appendix.

[3] The results of the tests by Hanck (2013) and Demetrescu et al. (2006) for the variables in first differences are omitted for brevity.

Table 5.2: Arsova and Örsal's (2019) intersection-type panel cointegration test

Trend in EC term				Trend orthogonal to EC term				Simes' crit. values	
Country	lag	$LR^{\mathrm{SL}}_{\mathrm{trace}}$	p-value	Country	lag	$LR^{\mathrm{SL}}_{\mathrm{trace}}$	p-value	5%	10%
				$H_0 : r = 0$					
Denmark	1	57.15	0.002**	Denmark	1	42.95	0.007	0.003	0.005
Poland	4	50.25	0.014	France	2	40.41	0.014	0.005	0.011
Sweden	2	49.50	0.017	Lithuania	2	39.00	0.021	0.008	0.016
France	2	44.62	0.059	Czech Republic	2	37.37	0.033	0.011	0.021
Lithuania	2	43.10	0.083	Estonia	4	34.92	0.062	0.013	0.026
Luxembourg	1	42.66	0.092	Sweden	2	34.08	0.077	0.016	0.032
Czech Republic	2	41.85	0.109	Portugal	1	33.36	0.091	0.018	0.037
Austria	2	41.77	0.111	Netherlands	2	32.45	0.111	0.021	0.042
Estonia	4	41.76	0.111	Germany	2	32.39	0.113	0.024	0.047
Germany	2	40.76	0.136	United Kingdom	2	31.62	0.133	0.026	0.053
United Kingdom	3	39.69	0.167	Spain	2	31.55	0.135	0.029	0.058
Norway	1	39.26	0.181	Belgium	3	30.72	0.161	0.032	0.063
Netherlands	2	39.23	0.183	Italy	2	30.43	0.171	0.034	0.068
Italy	2	38.86	0.195	Luxembourg	2	29.20	0.217	0.037	0.074
Switzerland	3	36.25	0.303	Finland	3	28.60	0.243	0.039	0.079
Portugal	1	35.13	0.358	Poland	4	28.52	0.247	0.042	0.084
Belgium	2	34.43	0.394	Austria	2	27.91	0.275	0.045	0.089
Spain	2	33.29	0.457	Switzerland	3	27.60	0.290	0.047	0.095
Finland	3	32.31	0.513	Norway	2	22.90	0.568	0.050	0.100
				$H_0 : r = 1$					
Czech Republic	2	27.73	0.063	Czech Republic	3	24.86	0.014	0.003	0.005
Netherlands	2	25.47	0.117	Lithuania	2	24.17	0.018	0.005	0.011
United Kingdom	2	24.62	0.146	United Kingdom	2	23.31	0.024	0.008	0.016
Spain	3	24.08	0.166	Luxembourg	1	18.94	0.093	0.011	0.021
Denmark	1	23.33	0.199	Switzerland	2	18.20	0.116	0.013	0.026
Sweden	2	23.00	0.214	Denmark	1	15.99	0.208	0.016	0.032
Germany	2	21.08	0.321	France	3	15.00	0.264	0.018	0.037
Italy	2	19.31	0.442	Germany	4	14.08	0.325	0.021	0.042
Switzerland	3	19.20	0.451	Spain	3	13.77	0.348	0.024	0.047
Luxembourg	3	18.53	0.501	Belgium	2	13.53	0.366	0.026	0.053
Estonia	4	17.79	0.558	Portugal	1	13.43	0.374	0.029	0.058
Portugal	1	16.69	0.643	Poland	4	13.37	0.378	0.032	0.063
Finland	4	15.75	0.714	Austria	3	12.49	0.450	0.034	0.068
Belgium	2	15.35	0.742	Finland	2	12.09	0.484	0.037	0.074
Austria	3	14.80	0.780	Italy	2	11.78	0.511	0.039	0.079
Poland	4	13.64	0.850	Netherlands	2	10.91	0.590	0.042	0.084
Norway	1	12.15	0.920	Estonia	2	9.06	0.754	0.045	0.089
Lithuania	4	11.72	0.935	Sweden	2	8.46	0.802	0.047	0.095
France	3	11.53	0.941	Norway	1	8.08	0.830	0.050	0.100

Notes: The lag order is selected according to the modified AIC criterion of Qu and Perron (2007). Results for each variable are sorted according to the p-values in ascending order for ease of comparison with the corresponding Simes' critical value.

*, ** and *** denote significance of the panel intersection test at the 10, 5 and 1% level, respectively.

common components (PANIC) is employed, as set out in Bai and Ng (2004). The time series are decomposed into unobserved common and idiosyncratic components and their integration and cointegration properties are analyzed separately. The benefit of such analysis is that it provides better understanding of the interconnections among the variables in the system.

Unobserved dynamic common factors are extracted by the method of principal components from the panel for each variable with country cross-sections. Prior to the extraction the observed data are first-differenced. For the panels of import prices, GDP and producer's costs they are also demeaned to account for the observed time trend. The data are also standardized to have unit variance. The number of unobserved common factors for each panel is selected by the criterion of Onatski (2010); the maximum number allowed is six. The criterion picks two factors for the panels of p and e, explaining 57% and 75% of the variation in the data, respectively. For each of the panels of y and w a single factor is chosen, explaining 52% and 23% of the total variation, respectively.

Once the estimated variable-specific common factors are extracted and subtracted from the first-differenced (and potentially demeaned) observations, the remaining residuals are accumulated to yield estimates $\hat{e}^x_{i,t}$ of the idiosyncratic components for each variable and cross-sectional unit, $x \in \{p, e, y, w\}$. The estimated idiosyncratic components are then tested for unit roots by the P_a, P_b, and $PMSB$ tests proposed by Bai and Ng (2010). Table 5.3 presents the results. The null hypothesis of a unit root cannot be rejected for either panel.

Table 5.3: Bai and Ng's (2010) panel unit root tests for the estimated idiosyncratic components

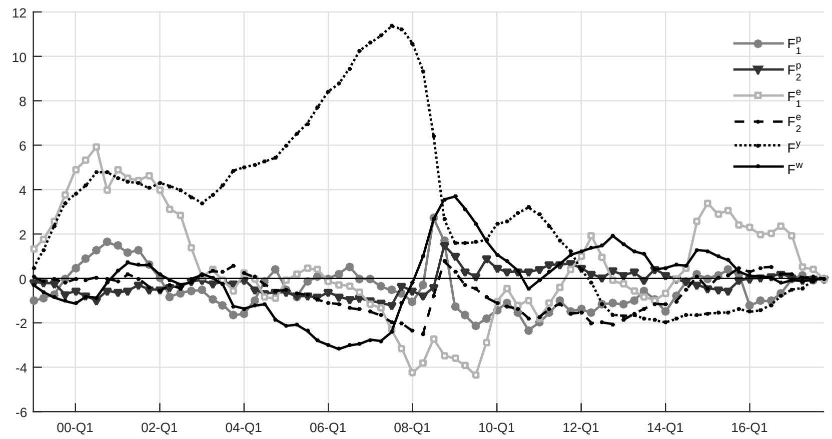
Idiosyncratic component	Avg. volatility	P_a	P_b	$PMSB$
$\hat{e}^p$	0.037	−0.005	−0.005	0.027
$\hat{e}^e$	0.058	1.221	1.825	2.893
$\hat{e}^y$	0.034	−0.758	−0.682	−0.585
$\hat{e}^w$	0.035	0.945	1.092	1.266

Notes: Trend is included in the test regressions for p, y and w, while only a constant is considered for e. All three test statistics have a $N(0,1)$ distribution under the null hypothesis of a unit root with a rejection region in the left tail of the distribution. The average volatility is computed over all cross-sections.

Next, the cointegration properties of the idiosyncratic components are examined by the PSL^J_{def} test of Arsova and Örsal (2018) and the $P^*_{\Phi^{-1}}$ test of Örsal and Arsova (2017). The first test computes the panel test statistic as the standardized average of the individual LR trace statistics of Saikkonen and Lütkepohl (2000) computed from defactored data, while the second one combines the p-values of these statistics by the inverse normal method. Both statistics have a limiting $N(0,1)$ distribution under the null hypothesis of a common cointegrating rank $H_0 : r_i = r, \forall i$, whereas the rejection region for the panel-SL test is in the right tail, and for the $P^*_{\Phi^{-1}}$ test in the left tail, respectively. Örsal and Arsova (2017) show that the $P^*_{\Phi^{-1}}$ exhibits better finite-sample properties than the panel-SL test in some situations. The value of the PSL^J_{def} test statistic under the null of no cointegration is 0.2, while that of the $P^*_{\Phi^{-1}}$ is -1.33, which is significant at the 10% level. As neither test rejects the null of cointegrating rank one ($PSL^J_{def} = -2.13$ and $P^*_{\Phi^{-1}} = 4.99$ in this case), we conclude that there is some, albeit not very strong, evidence of a single cointegrating relationship among the idiosyncratic components, matching the result for the observed variables.

We next turn our attention to the extracted and accumulated common factors. They are denoted as $F^p_1, F^p_2, F^e_1, F^e_2, F^y$, and F^w, with the superscript signifying the variable-specific panel they have been extracted from and the subscript denoting the factor number. A graph of the factors is displayed in Figure 5.1. It reveals how they all capture the effects of the Global Financial Crisis, reacting mostly simultaneously and with similar turns in the dynamics. Such behaviour hints at possible cointegration among them, which could lead to cross-unit cointegration of the observed variables.

Figure 5.1: Extracted common factors from the panel of each model variable



Testing for unit roots in the extracted common factors is carried out by a standard ADF test[4]. The results, presented in Table 5.4, indicate that the unit root null hypothesis cannot be rejected for any individual factor. The test statistic of the modified inverse normal panel test of Demetrescu et al. (2006) is -0.504, supporting this conclusion.

Table 5.4: Unit root tests for the estimated common factors

Factor	Volatility	deterministic term	lag order	ADF_τ	p-value
F_1^p	0.97	trend	2	-3.05	0.125
F_2^p	0.52	trend	4	-2.67	0.254
F^y	2.42	trend	1	-1.89	0.651
F^w	0.83	trend	2	-2.13	0.522
F_1^e	3.80	const	2	-2.17	0.496
F_2^e	1.53	const	1	-3.03	0.131

Notes: ADF_τ denotes the Augmented Dickey-Fuller test statistic. The lag order is selected according to the modified AIC (MAIC) criterion of Ng and Perron (2001). The p-values are computed as in MacKinnon (1996); the author is grateful to Christoph Hanck for providing the GAUSS code.

It is interesting to note the enormous difference between the volatility of the estimated common factors and that of the idiosyncratic components, displayed in Tables 5.3 and 5.4, respectively. Even the smallest volatility among those of the factors (0.52 for $\hat{F}_2^p$) is about ten times greater than the largest average volatility of the idiosyncratic components (0.058 for $\hat{e}^e$). Hence we conclude that it is the unobserved common factors

[4]Bai and Ng (2004) show that the limiting distributions of the ADF test statistics, computed for common factors extracted from first-differenced or first-differenced and demeaned data, coincide with the usual limiting distributions of the ADF test with a constant only or a constant and linear time trend, respectively.

which to a large extent determine the behaviour of the observed variables, while the idiosyncratic components have only a minor impact.

Having established the presence of global stochastic trends, we next assess whether they exhibit any cointegration. For more reliable results the SL test of Saikkonen and Lütkepohl (2000) is employed for each pair of estimated factors, as it is known that the LR cointegrating rank tests become less powerful in larger systems (see, e.g. Saikkonen and Lütkepohl (2000)). Table 5.5 displays the results.

Table 5.5: SL cointegration tests for the estimated common factors

Trend in EC term				Trend orthogonal to EC term			
Factors	Lag order	$LR^{\text{SL}}_{\text{trace}}$	p-value	Factors	Lag order	$LR^{\text{SL}}_{\text{trace}}$	p-value
F_1^p, F_2^p	2	20.81	0.006***	F_2^p, F^w	4	19.59	0.001***
F_2^p, F^w	4	18.01	0.020**	F_1^p, F^w	3	16.81	0.002***
F_1^p, F^w	3	16.90	0.032**	F_1^p, F_2^p	2	16.60	0.003***
F_2^e, F^w	3	13.75	0.105	F^y, F^w	4	14.53	0.007***
F_1^e, F^y	3	12.52	0.159	F_2^e, F^w	3	13.69	0.010***
F_2^p, F_1^e	4	12.45	0.163	F_1^e, F^w	2	12.88	0.014**
F_1^e, F^w	2	12.43	0.163	F_1^e, F^y	3	12.50	0.016**
F_1^p, F_1^e	3	11.73	0.205	F_1^P, F^y	3	11.91	0.021**
F_2^p, F^y	4	11.15	0.244	F_2^p, F_1^e	4	11.59	0.024**
F_1^p, F_2^e	2	10.63	0.284	F_2^p, F^y	4	11.45	0.026**
F_1^P, F^y	3	10.12	0.327	F_1^p, F_2^e	2	10.82	0.034**
F^y, F^w	4	9.96	0.342	F_1^p, F_1^e	3	10.03	0.048**
F_2^e, F^y	3	7.76	0.574	F_2^e, F^y	3	7.78	0.120
F_2^p, F_2^e	4	5.99	0.776	F_2^p, F_2^e	4	6.34	0.208
F_1^e, F_2^e	1	5.14	0.858	F_1^e, F_2^e	1	5.50	0.282

Notes: The lag order is selected according to the modified AIC criterion of Qu and Perron (2007). Results for each variable are sorted according to the p-values in ascending order for ease of comparison with the corresponding critical value of Hommel's (1988) procedure.
*, ** and *** denote significance at the 10, 5 and 1% level, respectively.

At first glance there seems to exist a cointegrating relationship between almost any pair of factors considered when allowing for no trend in the cointegrating relation. However, the results of these tests are highly correlated, and one must take the nature of such multiple testing into account. In order to select only the meaningful rejections, Hommel (1988) proposes a procedure which controls the family-wise error rate at a chosen significance level α. Details on Hommel's procedure can be found in Hanck (2013), whose exposition is briefly reproduced here for convenience. Let the ordered p-values of n tests be $p^*_{(1)} \leq \ldots \leq p^*_{(n)}$ and $\mathbb{N}_n$ denote the set of all natural numbers between 1 and n. Selecting the meaningful rejections by the Hommel's procedure is then carried out in two steps: (A) Compute $j = \max\{i \in \mathbb{N}_n : p^*_{(n-i+k)} > \frac{k\alpha}{i}, \forall k \in \mathbb{N}_i\}$, and (B) If $p^*_{(n)} \leq \alpha$, reject all $H_{i,0}$; else, reject those $H_{i,0}$ for which $p^*_i \leq \frac{\alpha}{j}$.

Following this procedure, $j = 10$ is computed, and the corresponding Hommel's critical values at the 5% and 10% significance levels are 0.005 and 0.01, respectively. Hence, only the first five rows in the second panel of Table 5.5 can be considered genuine rejections at the 10%-level; at the 5%-level it would only be the first three. We may therefore conclude that two global stochastic trends exist among the extracted common factors: one which is shared by F_1^p, F_2^p, F^y, F^w and F_2^e, and one driving F_1^e.

By analyzing the factor loadings (see Table 5.10 in the Appendix), F_1^e may be identified as the Euro-exchange-rate factor, which is perhaps not surprising, as the dataset features both countries in and outside the euro area. On the other hand, F_2^e can be thought of the factor influencing more the dynamics of the exchange rates of the non-euro area countries (including the newest members of the euro area like Lithuania, for example). Relating these results to those from the cointegration testing of the observed variables (Table 5.2), we conclude that there is much more evidence in favour of a long-run equilibrium relationship in the ERPT for non-euro area countries than it is for euro area ones. One explanation for this phenomenon may lie in the fact that the import prices in euro area countries, whose principal share of imports come from other euro area countries, react much less to aggregate exchange rate fluctuations because these are basically zero between the one and the same currency. This leads us to believe that the ERPT estimates would be lower for the older member-countries of the euro area than they would be for the newer ones or the countries outside the euro area.

The results of the unit root and cointegration analysis can be summarized as follows. All variables in the log-linear ERPT relationship in eq. (5.6) are integrated of order one. There is evidence of a single cointegrating relationship at the panel level linking the observed variables, suggesting that the average long-run elasticity of the exchange rate is different from zero. It is worth noting that, contrary to the results of De Bandt et al. (2008), this relationship emerges without the necessity to consider structural breaks, neither in the deterministic components, nor in the long-run equilibrium. This is so because of the present cross-unit cointegration driven by unobserved common factors. These factors capture the major exogenous shocks such as the Global Financial Crisis which, in turn, force the observed variables to react more or less simultaneously and in a similar fashion. Although the data dynamics are mostly determined by six unobserved common factors (two for the panel of import prices, two for the panel of nominal exchange rate and one for each of the domestic demand and producer's cost proxy panels), the driving forces behind them are only two distinct global stochastic trends. One of them is shared by the import prices panel, the domestic demand panel, the producer's costs panel, and by the exchange rate data for countries outside the euro area as well as newer member-countries of the euro area. The second global stochastic trend can be viewed as a Euro-nominal-exchange-rate factor, influencing mostly the exchange rate series of the euro area countries. The idiosyncratic components of the data, although with much less impact than the common components, are also non-stationary and cointegrated by a single relationship. These findings lead us to expect more significant ERPT elasticities for non-euro area countries than for euro area ones.

5.4 ERPT estimation

Having established the presence of a long-run equilibrium relationship at the panel level, the next step is to estimate the ERPT equation (5.6). However, the presence of cross-sectional dependence, depending on its nature, may yield the results of earlier panel regression estimators either biased, inconsistent or inefficient (see, e.g. Phillips and Sul (2003, 2007) and Moon and Weidner (2017)). Further, cross-unit cointegration has also been shown by Urbain and Westerlund (2006) to pose an issue in pooled ordinary least squares estimation. Hence an estimator which takes into account both cointegration and cross-sectional dependence induced by global stochastic trends is needed.

Two suitable approaches have recently been proposed in the literature. The first one, put forward by Bai et al. (2009), features two estimators: the continuously-updated bias-corrected (Cup-BC) and the continuously-updated fully-modified (Cup-FM) estimator. They estimate level relationships in panel cointegration models where unobserved common factors drive the dependence in the regression errors and which may also be correlated with the regressors. This methodology has been widely applied in the recent empirical panel data studies, employed by e.g. Bodart et al. (2015) for estimation of a long-run relationship between real exchange rates and commodity prices, and by Örsal (2017) for estimation of a long-run money demand relation.

The second approach, using a common correlated effects (CCE) mean-group (MG) estimator, is due to Chudik and Pesaran (2015). They extend earlier work of Pesaran (2006) to panel data models allowing for lagged dependent variables and weakly exogenous regressors. The residual dependence induced by the unobserved common factors is captured by cross-sectional averages of the observed variables included as additional regressors in the individual equations. Details on the estimation by each estimator are briefly outlined next, while the empirical results are discussed in Section 5.4.3.

5.4.1 The Cup-BC and Cup-FM estimators of Bai et al. (2009)

The ERPT equation (5.6) can be written in the Bai et al. (2009) estimation framework as

$$p_{it} = \beta_0 + \beta_1 e_{it} + \beta_2 y_{it} + \beta_3 w_{it} + u_{it}, \tag{5.8}$$

$$u_{it} = \lambda_i' f_t + \varepsilon_{it}. \tag{5.9}$$

The errors ε_{it} are assumed to be stationary and only weakly cross-sectionally dependent, while the unobserved common factors in the $(r \times 1)$-vector f_t are allowed to be $I(0)$, $I(1)$ or a mixture of the two. They are treated as parameters and estimated together with the common slope coefficients $\beta = (\beta_1, \beta_2, \beta_2)$ in an iterative procedure. Albeit consistent, the resulting $\hat{\beta}_{\text{Cup}}$ estimator has been shown to be asymptotically biased, hence a bias correction is necessary. The Cup-BC and the Cup-FM estimators differ with regard to when this bias correction takes place. With the Cup-BC it is applied only once at the final iteration, while with the Cup-FM the correction is made at each iteration. $\hat{\beta}_{\text{Cup}}$ is shown to be at least T-consistent regardless of the integration order (zero or one) of the factors or that of the regressors. Being pure panel estimators, however, both the Cup-FM and the Cup-BC assume homogeneity of the coefficients across cross-sections, and hence do not produce individual-unit results, which may be viewed as a drawback in practice.

In order to account for the trending behaviour of the variables p, y and w, eq. (5.9) is estimated with demeaned and detrended series, as suggested by Bai et al. (2009). The number r of residual common factors is selected by the criterion of Onatski (2010). It picks two factors which account for 59% of the variance of the first-stage residuals $\hat{u}_{it}$.

The results are presented in Table 5.6. The actual estimates of the elasticity parameters are quite similar across the two estimators. The nominal exchange rate elasticity, which is the only statistically significant coefficient, is estimated by both the Cup-FM and the Cup-BC as $\hat{\beta}_1 = -0.37$. This implies that a 1% depreciation in the exchange rate would lead to an average increase of 0.37% in the import prices. The results are discussed in more detail in Section 5.4.3.

Table 5.6: ERPT estimation results by the Cup-BC and Cup-FM estimators

Variable	Cup-BC	Cup-FM
Nominal exchange rate elasticity $\hat{\beta}_1$	-0.372^{***}	-0.369^{***}
	[0.029]	[0.029]
Domestic demand elasticity $\hat{\beta}_2$	-0.041	-0.009
	[0.044]	[0.042]
Producers' costs elasticity $\hat{\beta}_3$	-0.011	0.007
	[0.035]	[0.035]

Notes: Standard errors are presented in brackets.
*, ** and *** denote significance at the 10, 5 and 1% level, respectively.

Figure 5.2 in the Appendix presents a graph of the estimated residual common factors[5]. The Global Financial Crisis manifests itself in the two spikes in 2008Q4 and 2009Q1, respectively. Analysis of the Cup-FM[6] model residuals, depicted in Figure 5.3 in the Appendix, reveals that the two factors adequately capture the effects of the crisis, as no further common shocks can be observed. Applying Demetrescu et al.'s (2006) panel unit root test to the estimated residuals yields a value of -5.25 for Hartung's test

[5]These factors, common to the first-stage residuals of the Cup-FM and Cup-BC models, are not to be confused with the common factors extracted from the panel formed by each variable.

[6]Results for the residuals of the CUP-BC estimation are very similar and omitted for brevity.

statistic with $\kappa = 0.2$, which points to their stationarity. This leads us to the conclusion that no model assumptions have been violated.

5.4.2 The dynamic CCE estimator of Chudik and Pesaran (2015)

The second approach considered for the estimation of the ERPT equation (5.6) is the panel autoregressive distributed-lag (ARDL) model with multifactor error structure proposed by Chudik and Pesaran (2015). This framework differs from the specification of Bai et al. (2009) in that it allows for (a) lagged values of the dependent and independent variables as additional regressors and (b) heterogeneous coefficients $\beta_{j,i}, (i = 1, \dots, N,\ j = 0, \dots, 3)$, for each unit, which are then combined in a mean-group (MG) estimator. As in Bai's framework, unobserved common factors in the residuals drive the strong cross-sectional dependence. The common factors, however, are not explicitly estimated from the data. Instead, they are approximated by cross-sectional averages of the observed model variables and the resulting regressions are estimated individually for each unit by ordinary least squares. A necessary condition for the validity of the resulting CCE MG estimator is that the number of unobserved common factors be no more than the observed variables in the system. This assumption is likely to be satisfied in our case, as Onatski's (2010) criterion picks two factors in the residuals of the panel regression in eq. (5.9). Initially proposed for stationary factors (Pesaran, 2006), the CCE approach has been proved to be valid for integrated factors as well (Kapetanios et al., 2011). With regard to the assumed weak exogeneity of the regressors, also necessary for the validity of the DCCE MG estimator, we note that the preceding analysis is valid upon the assumption that changes in the import prices do not contemporaneously affect exchange rates, domestic demand or producers' costs. This assumption is commonly made in the empirical ERPT literature and is not a restrictive one, given the quarterly frequency of the data; for a more detailed discussion we refer to Brun-Aguerre et al. (2017).

To cast the ERPT model (5.6) into the CCE-framework, we begin with eq. (5.9), allowing for heterogeneous coefficients[7]:

$$p_{it} = \beta_{0,i} + \beta_{1,i} e_{it} + \beta_{2,i} y_{it} + \beta_{3,i} w_{it} + u_{it}, \tag{5.10}$$

$$u_{it} = \lambda_i' f_t + \varepsilon_{it}. \tag{5.11}$$

Taking the cointegrating relationship explicitly into account, we then put it in an

[7]The exposition is similar to that of Eberhardt and Presbitero (2015), whose Stata code for the dynamic CCE MG estimator has been used for the estimation.

error-correction form:

$$\Delta p_{it} = \beta_{0,i} + \rho_i \left(p_{i,t-1} - \beta_{1,i}e_{i,t-1} - \beta_{2,i}y_{i,t-1} - \beta_{3,i-1}w_{it} - \lambda_i' f_{t-1}\right) \\ + \gamma_{1,i}\Delta e_{it} + \gamma_{2,i}\Delta y_{it} + \gamma_{3,i}\Delta w_{it} + \gamma_{i,f}\Delta f_t + \varepsilon_{it}, \tag{5.12}$$

so that, by re-arranging, we get

$$\Delta p_{it} = \pi_{0,i} + \pi_{ec,i}p_{i,t-1} + \pi_{1,i}e_{i,t-1} + \pi_{2,i}y_{i,t-1} + \pi_{3,i}w_{i,t-1} + \pi_{1f,i}f_{t-1} \\ + \pi_{4,i}\Delta e_{it} + \pi_{5,i}\Delta y_{it} + \pi_{6,i}\Delta w_{it} + \pi_{2f,i}\Delta f_t + \varepsilon_{it}. \tag{5.13}$$

The long-run parameters $\beta_{j,i}, (i = 1, \ldots, N,\ j = 1, 2, 3)$, can be recovered from the coefficients of eq. (5.13) as $\beta_{j,i} = -\pi_{j,i}/\pi_{ec,i}$, while the short-run parameters $\pi_{j,i}, j = 4, 5, 6$, are estimated directly. The term $\pi_{ec,i}$ describes the speed of adjustment to equilibrium, and its statistical significance may be viewed as an additional evidence of the presence of cointegration.

For the estimation, the unobserved common factors f_t in eq. (5.13) are replaced by cross-sectional averages of the observed variables:

$$\Delta p_{it} = \pi_{0,i} + \pi_{ec,i}p_{i,t-1} + \pi_{1,i}e_{i,t-1} + \pi_{2,i}y_{i,t-1} + \pi_{3,i}w_{i,t-1} \\ + \pi_{4,i}\Delta e_{it} + \pi_{5,i}\Delta y_{it} + \pi_{6,i}\Delta w_{it} \\ + \pi^*_{1,i}\Delta \bar{p}_t + \pi^*_{2,i}\bar{p}_{t-1} + \pi^*_{3,i}\bar{e}_{t-1} + \pi^*_{4,i}\bar{y}_{t-1} + \pi^*_{5,i}\bar{w}_{t-1} \\ + \pi^*_{6,i}\Delta \bar{e}_t + \pi^*_{7,i}\Delta \bar{y}_t + \pi^*_{8,i}\Delta \bar{w}_t + \varepsilon_{it}. \tag{5.14}$$

So far, eq. (5.14) constitutes the model for the standard CCE MG estimator of Pesaran (2006). As Chudik and Pesaran (2015) show, finite-sample bias arises in the dynamic panel model with weakly exogenous regressors and recommend the inclusion of sufficient number (s) lagged values of the cross-sectional averages to mitigate it. Their suggested rule of thumb, $s = \text{int}(T^{1/3})$, gives $\hat{s} = 4$ in our case. Hence the complete model to be estimated by the dynamic CCE (DCCE) estimator becomes

$$\Delta p_{it} = \pi_{0,i} + \pi_{ec,i}p_{i,t-1} + \pi_{1,i}e_{i,t-1} + \pi_{2,i}y_{i,t-1} + \pi_{3,i}w_{i,t-1} \\ + \pi_{4,i}\Delta e_{it} + \pi_{5,i}\Delta y_{it} + \pi_{6,i}\Delta w_{it} \\ + \pi^*_{1,i}\Delta \bar{p}_t + \pi^*_{2,i}\bar{p}_{t-1} + \pi^*_{3,i}\bar{e}_{t-1} + \pi^*_{4,i}\bar{y}_{t-1} + \pi^*_{5,i}\bar{w}_{t-1} \\ + \pi^*_{6,i}\Delta \bar{e}_t + \pi^*_{7,i}\Delta \bar{y}_t + \pi^*_{8,i}\Delta \bar{w}_t \\ + \sum_{l=1}^{4} \pi^*_{9,l,i}\Delta \bar{p}_{t-l} + \sum_{l=1}^{4} \pi^*_{10,l,i}\Delta \bar{e}_{t-l} \\ + \sum_{l=1}^{4} \pi^*_{11,l,i}\Delta \bar{y}_{t-l} + \sum_{l=1}^{4} \pi^*_{12,l,i}\Delta \bar{w}_{t-l} + \varepsilon_{it}. \tag{5.15}$$

The results from the estimation at the panel level are listed in Table 5.7, while results from the individual country models are available in Table 5.11 in the Appendix.

Table 5.7: ERPT estimation results by the CCE MG estimator

Variable		Standard MG estimator const	trend	DCCE estimator const	trend	Augmented DCCE estimator const	trend
e	LRA	−0.473***	−0.479***	−0.420***	−0.315***	−0.441***	−0.325***
		[0.069]	0.082	[0.121]	[0.095]	[0.154]	[0.131]
	SRA	−0.363***	−0.370***	−0.305***	−0.221**	−0.240***	−0.187*
		[0.073]	0.062	[0.071]	[0.104]	[0.081]	[0.106]
y	LRA	0.473***	0.337***	−0.063	0.083	−0.003	0.112
		[0.089]	0.089	[0.115]	[0.170]	[0.117]	[0.158]
	SRA	−0.074	−0.163	0.000	0.105	0.069	0.106
		[0.138]	0.137	[0.131]	[0.131]	[0.091]	[0.116]
w	LRA	−0.124	−0.163***	0.026	0.132*	0.031	0.056
		[0.080]	0.044	[0.065]	[0.072]	[0.092]	[0.094]
	SRA	0.116**	0.092*	0.022	0.028	0.029	0.024
		[0.047]	0.049	[0.070]	[0.071]	[0.067]	[0.068]
avg. EC coefficient $\bar{\rho}$		−0.363***	−0.428***	−0.650***	−0.729***	−0.666***	−0.748***
		[0.035]	0.035	[0.065]	[0.064]	[0.052]	[0.045]
Implied half-life		1.54	1.24	0.66	0.53	0.63	0.50
RMSE		0.022	0.021	0.012	0.012	0.010	0.010
CD test		34.87	34.13	4.50	3.12	1.90	0.52
Trends share			0.47		0.42		0.37

Notes: Standard MG estimator refers to the MG estimator of Pesaran and Smith (1995). Augmented DCCE estimator denotes the DCCE estimator augmented with a dummy variable accounting for the Great Recession in 2008. LRA and SRA denote long-run average and short-run average coefficients, respectively. Standard errors are presented in brackets. Implied half-life is computed as $\ln(0.5)/\ln(1+\bar{\rho})$. CD test denotes Pesaran's (2015) test for weak cross-sectional dependence of the regression residuals; it has $N(0,1)$ distribution under the null hypothesis. Trends share stands for the share of group-specific trends significant at the 5% level.
*, ** and *** denote significance at the 10, 5 and 1% level, respectively.

The first two columns present the results from the estimation with the standard mean-group panel estimator of Pesaran and Smith (1995), included to illustrate what the effect of unattended cross-secitonal dependence would be. The residual correlation as measured by Pesaran's (2015) CD test is highly significant and the resulting estimates are therefore likely to be biased and inconsistent (Moon and Weidner, 2017).

The middle two columns of Table 5.7 present the results from the DCCE estimator with a constant or a constant and a linear time trend, respectively. The CD test statistic has much lower values, indicating that the inclusion of the cross-sectional averages controls well for the dependence. Nevertheless, it is significant at the 1% level. Analysis of the regression residuals has identified a common shock due to the Great Recession as a possible reason for the elevated correlation. Hence the model equation (5.15) is augmented by a dummy variable which has the value one in 2008Q3, 2008Q4 and 2009Q1, and zero otherwise. It is then included as an observed common factor in the estimation.

The results from this augmented DCCE model are presented in the last two columns of Table 5.7. The CD test statistic (1.90) is significant at the 10% level in the constant only case, showing that the residual cross-sectional correlation has not been eliminated completely even after the inclusion of the Great Recession dummy variable. Such correlation might, however, be due to omitted incidental trends. Turning to estimation results with trend included (last column), we note that this is indeed the case. The value of the CD statistic is 0.52, meaning that there is no strong cross-sectional dependence left among the residuals. Also more than a third (37%) of the time trend coefficients are significant at the 5% level, suggesting that the trending behaviour of the variables must be accounted for. The insignificant value of the CD test statistic serves also as evidence that the number of unobserved common factors is less than the number of observed variables in the system, thus rendering the DCCE estimator valid in this regard.

The residuals of the augmented DCCE model with trend are analyzed as a diagnostic check. Demetrescu et al.'s (2006) panel unit root test yields a value of -17.64 for Hartung's inverse normal test with $\kappa = 0.2$, convincingly rejecting the unit root null hypothesis. A plot of the estimated residuals, presented in Figure 5.4 in the Appendix, reveals no anomalies which could have resulted from potentially unattended structural breaks.

The next section compares the results from the Bai et al.'s (2009) estimators with those from the augmented DCCE estimator with trend and discusses their implications.

5.4.3 Discussion

The results produced by the Cup-BC, Cup-FM and the DCCE estimators are remarkably similar. The point estimate of the ERPT by both the Cup-BC and Cup-FM estimators is -0.37 with a standard error of 0.03, while the average long-run ERPT by the DCCE MG estimator is -0.33 with a standard error of 0.13. Hence their 95% confidence bands overlap, and neither includes the borderline values zero or one. Therefore, there is no evidence of PCP or LCP in the long-run, but rather of incomplete and low ERPT at the panel level. The average short-run ERPT coefficient estimated by the DCCE model is -0.19, which is significantly different from zero at the 10% level, but not at the 5%. These values are lower than the average exchange rate elasticity of 0.54 reported by Ben Cheikh and Rault (2016) for twelve euro-area countries in the period 1990Q3 - 2012Q4. Thus our results provide further evidence that ERPT has been declining over time, as found in the recent literature (Campa and Goldberg, 2005, Ben Cheikh and Rault, 2016). This may be attributed to the fact that our data comprises a longer period since the creation of the monetary union, so that a greater degree of convergence to more stable macroeconomic conditions has taken place in most countries of the panel. Another reason, however, might be the significant share of intra-euro area trade, which biases the aggregate ERPT estimates downwards (see Blagov, 2018).

Turning our attention to the average error-correction coefficient $\bar{\rho}$, reported in the last column of Table 5.7, we see that it is highly significant, once again highlighting the presence of cointegration in the system. Its value is -0.748, which implies high speed of adjustment to equilibrium – the aggregate implied half-life[8] is just 0.5 quarters.

The importing country's demand and the producers costs proxy are not statistically significant in either model. Domestic demand y_{it} not being an informative regressor for import prices despite predictions by economic theory is a result found also by other empirical studies, see, e.g., Campa and Goldberg (2005) and Beirne and Bijsterbosch (2009). The insignificance of the producers costs elasticity, on the other hand, could be based on unit labour costs not being a sufficiently good approximation for the producers' actual costs due to their slowly changing nature, for example. In a single-equation framework such results could raise questions regarding the validity of the ERPT results, as the estimation might be subject to omitted variable bias. In the panel framework, however, the unobserved common factors become a remedy for such problem. Being allowed to be $I(0)$, $I(1)$ or mixture of the two by both models, they capture the effects of any common shocks that influence the import prices and are not contained in the regressors. Hence besides controlling for cross-sectional dependence, the inclusion of unobserved common factors in the regression equations (5.9) and (5.13) has the added benefit of guarding against omitted variable bias. Therefore, the results for the exchange rate elasticity remain valid.

Finally, we discuss the individual countries' results from the DCCE estimator, presented in Table 5.11 in the Appendix. The estimated long-run exchange rate coefficients are quite heterogeneous with relatively large standard errors, which is the price to pay for the inclusion of current and lagged values of the cross-sectional averages in the individual equations. The long-run exchange rate elasticity is insignificantly different from zero for most of the older euro area member countries (Austria, Belgium, Finland, France, Luxembourg, Netherlands, Spain), and also for Switzerland and Estonia. Notable exceptions are Germany, Italy, and Portugal. The estimated long-run pass-through coefficients for these countries are -0.51, -1.08 and -1.26, respectively. On the other hand, the long-run ERPT coefficients are significant and with the expected negative sign for all non-euro area countries except for Switzerland: that is, for the Czech Republic (-0.37), Denmark (-1.46), Lithuania (-1.17), Poland (-0.30), Sweden (-0.82) and the United Kingdom (-0.47).

5.5 Summary and outlook

This paper establishes the presence of a long-run equilibrium relationship underlying the ERPT in a panel of nineteen European countries by employing new second-generation panel cointegration tests. Taking into account that unobserved global stochastic trends

[8]The period of time needed for deviations in import prices to decline by half following a unit shock of the exchange rate.

drive the cointegrating relationship and induce cross-sectional dependence in the data, aggregate long- and short-run exchange rate elasticities are estimated by novel panel estimators from most recent data. The results support the findings of earlier studies which report declining ERPT over time.

Future work could broaden the scope of the present analysis in several directions. Firstly, other data on import prices, which distinguish between imports from inside and outside the euro area, could be considered. As argued in Section 3, since a significant share of a Eurozone country's imports comes from other Eurozone contires, this may introduce a downward bias in an aggregate ERPT estimate. Monthly data on intra- and extra-euro area import prices is available from Eurostat, however sufficiently long time series exist for only six countries and the euro area as a whole. Hence the large-N panel estimators in this study would not be applicable and other estimation techniques would be needed. Secondly, ERPT to import prices could be investigated at an industry level, as considerable heterogeneity in ERPT elasticities has been reported across different industries (De Bandt et al., 2008; Blagov, 2018). Finally, the question of asymmetric ERPT as in Brun-Aguerre et al. (2017) seems a promising avenue to explore, in particular with regard to providing a better understanding of the pricing behaviour of exporters w.r.t. currency appreciation or depreciation and its implications for policy makers.

Bibliography

Arsova, A., Örsal, D.D.K., 2018. Likelihood-based panel cointegration test in the presence of a linear time trend and cross-sectional dependence. Econometric Reviews 37, 1033–1050.

Arsova, A., Örsal, D.D.K., 2019. Intersection tests for the cointegrating rank in dependent panel data. Forthcoming in Communications in Statistics: Simulation and Computation , doi:10.1080/03610918.2018.1489552.

Bai, J., Kao, C., Ng, S., 2009. Panel cointegration with global stochastic trends. Journal of Econometrics 149, 82–99.

Bai, J., Ng, S., 2004. A PANIC Attack on Unit Roots and Cointegration. Econometrica 72, 1127–1177.

Bai, J., Ng, S., 2010. Panel unit root tests with cross-section dependence: a further investigation. Econometric Theory 26, 1088–1114.

Bailliu, J., Fujii, E., 2004. Exchange rate pass-through and the inflation environment in industrialized countries: an empirical investigation. Working Paper 2004-21. Bank of Canada.

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

Beirne, J., Bijsterbosch, M., 2009. Exchange rate pass through in Central and Eastern European member states. Working Paper 1120. European Central Bank.

Ben Cheikh, N., Rault, C., 2016. Recent estimates of exchange rate pass-through to import prices in the euro area. Review of World Economics 152, 69–105.

Ben Cheikh, N., Rault, C., 2017. Investigating first-stage exchange rate pass-through: Sectoral and macro evidence from euro area countries. The World Economy 40, 2611–2638.

Blagov, B., 2018. Exchange rate uncertainty and import prices in the euro area. Ruhr Economic Papers No. 789. RWI - Leibniz Institute for Economic Research.

Bodart, V., Candelon, B., Carpantier, J.F., 2015. Real exchanges rates, commodity prices and structural factors in developing countries. Journal of International Money and Finance 51, 264–284.

Brun-Aguerre, R., Fuertes, A.M., Greenwood-Nimmo, M., 2017. Heads I win; tails you lose: asymmetry in exchange rate pass-through into import prices. Journal of the Royal Statistical Society: Series A (Statistics in Society) 180, 587–612.

Brun-Aguerre, R., Fuertes, A.M., Phylaktis, K., 2012. Exchange rate pass-through into import prices revisited: What drives it? Journal of International Money and Finance 31, 818–844.

Campa, J.M., Goldberg, L.S., 2005. Exchange Rate Pass-Through into Import Prices. Review of Economics and Statistics 87, 679–690.

Chudik, A., Pesaran, M.H., 2015. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. Journal of Econometrics 188, 393–420.

Chudik, A., Pesaran, M.H., Tosetti, E., 2011. Weak and strong cross-section dependence and estimation of large panels. The Econometrics Journal 14, C45–C90.

De Bandt, O., Banerjee, A., Kozluk, T., 2008. Measuring Long-Run Exchange Rate Pass-Through. Economics: The Open-Access, Open-Assessment E-Journal 2, 1.

De Bandt, O., Razafindrabe, T., 2014. Exchange rate pass-through to import prices in the Euro-area: A multi-currency investigation. International Economics 138, 63–77.

Delatte, A.L., López-Villavicencio, A., 2012. Asymmetric exchange rate pass-through: Evidence from major countries. Journal of Macroeconomics 34, 833–844.

Demetrescu, M., Hassler, U., Tarcolea, A.I., 2006. Combining Significance of Correlated Statistics with Application to Panel Data. Oxford Bulletin of Economics and Statistics 68, 647–663.

Eberhardt, M., Presbitero, A.F., 2015. Public debt and growth: Heterogeneity and non-linearity. Journal of International Economics 97, 45–58.

Hanck, C., 2013. An Intersection Test for Panel Unit Roots. Econometric Reviews 32, 183–203.

Hartung, J., 1999. A Note on Combining Dependent Tests of Significance. Biometrical Journal 41, 849–855.

Hommel, G., 1988. A stagewise rejective multiple test procedure based on a modified Bonferroni test. Biometrika 75, 383–386.

Kapetanios, G., Pesaran, M.H., Yamagata, T., 2011. Panels with non-stationary multifactor error structures. Journal of Econometrics 160, 326–348.

MacKinnon, J.G., 1996. Numerical distribution functions for unit root and cointegration tests. Journal of Applied Econometrics 11, 601–618.

Moon, H.R., Weidner, M., 2017. Dynamic linear panel regression models with interactive fixed effects. Econometric Theory 33, 158–195.

Ng, S., Perron, P., 2001. LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power. Econometrica 69, 1519–1554.

Onatski, A., 2010. Determining the Number of Factors from Empirical Distribution of Eigenvalues. Review of Economics and Statistics 92, 1004–1016.

Örsal, D.D.K., 2017. Analysing money demand relation for OECD countries using common factors. Applied Economics 49, 6003–6013.

Örsal, D.D.K., Arsova, A., 2017. Meta-analytic cointegrating rank tests for dependent panels. Econometrics and Statistics 2, 61–72.

Pesaran, M., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. Journal of Econometrics 68, 79–113.

Pesaran, M.H., 2006. Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. Econometrica 74, 967–1012.

Pesaran, M.H., 2007. A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22, 265–312.

Pesaran, M.H., 2015. Testing Weak Cross-Sectional Dependence in Large Panels. Econometric Reviews 34, 1089–1117.

Phillips, P.C., Sul, D., 2007. Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. Journal of Econometrics 137, 162–188.

Phillips, P.C.B., Sul, D., 2003. Dynamic panel estimation and homogeneity testing under cross section dependence. The Econometrics Journal 6, 217–259.

Qu, Z., Perron, P., 2007. A Modified Information Criterion for Cointegration Tests Based on a VAR Approximation. Econometric Theory 23, 638–685.

Saikkonen, P., Lütkepohl, H., 2000. Trend Adjustment Prior to Testing for the Cointegrating Rank of a Vector Autoregressive Process. Journal of Time Series Analysis 21, 435–456.

Simes, R.J., 1986. An Improved Bonferroni Procedure for Multiple Tests of Significance. Biometrika 73, 751.

Urbain, J.P., Westerlund, J., 2006. Spurious Regression in Nonstationary Panels with Cross-Unit Cointegration. Working Paper RM/06/057. Universiteit Maastricht.

5.A Appendix

5.A.1 Simes' (1986) intersection test

Let $p_1^* \leq \ldots \leq p_N^*$ be the p-values of N individual test statistics. Ordering them as $p_{(1)}^* \leq \ldots \leq p_{(N)}^*$, the joint null hypothesis $H_0 := \bigcap_i H_{i,0},\ i = 1, \ldots, N$, (that all individual null hypotheses are simultaneously true) is rejected by Simes' test at significance level α if

$$p_{(i)}^* \leq \frac{i\alpha}{N} \text{ for any } i = 1, \ldots, N. \tag{5.16}$$

Simes (1986) shows that the test is conservative under independence of the individual test statistics, that is

$$P_{H_0}\left\{p_{(i)}^* \geq \frac{i\alpha}{N}, i = 1, \ldots, N\right\} \geq 1 - \alpha. \tag{5.17}$$

Simes' intersection method has been introduced to testing for panel unit roots by Hanck (2013) and to testing for panel cointegration in dependent panels by Arsova and Örsal (2019).

5.A.2 Demetrescu et al.'s (2006) panel unit root test

Demetrescu et al. (2006) employ the modified inverse normal method of Hartung (1999), in which the dependence is captured by a single correlation coefficient ρ_t, which can be interpreted as a "mean correlation approximating the case of possibly different correlations between the transformed statistics" (Hartung, 1999). The statistic has a $N(0,1)$ distribution under the null hypothesis. It is computed as [9]

$$t\left(\hat{\rho}_{\mathrm{t}}^*, \kappa\right) = \frac{\sum_{i=1}^N t_i}{\sqrt{N + (N^2 - N)\left(\hat{\rho}_{\mathrm{t}}^* + \kappa \cdot \sqrt{\frac{2}{(N+1)}}(1 - \hat{\rho}_{\mathrm{t}}^*)\right)}}. \tag{5.18}$$

Here $t_i = \Phi^{-1}(p_i^*)$ denote the probits and the variance of the denominator is augmented with an estimator of the correlation between the individual probits $\hat{\rho}_{\mathrm{t}}^*$: $\hat{\rho}_{\mathrm{t}}^* = \max\{-\frac{1}{N-1}, \hat{\rho}_{\mathrm{t}}\}$, where $\hat{\rho}_{\mathrm{t}} = 1 - \frac{1}{N-1}\sum_{i=1}^N \left(t_i - \frac{1}{N}\sum_{i=1}^N t_i\right)^2$.

The correction term $\kappa\sqrt{\frac{2}{(N+1)}}(1 - \hat{\rho}_{\mathrm{t}}^*)$, which simply scales the standard deviation of $\hat{\rho}_{\mathrm{t}}$ by a factor κ, aims to avoid a systematic underestimation of the denominator in eq. (5.18). For the κ parameter Hartung suggests two alternative values: $\kappa_1 = 0.2$ and $\kappa_2 = 0.1 \cdot \left(1 + \frac{1}{N-1} - \hat{\rho}_{\mathrm{t}}^*\right)$, where κ_2 is suitable mainly for smaller $\hat{\rho}_{\mathrm{t}}^*$.

[9] For simplicity the computation of the statistic is presented with unit weights for all cross-sections, as in its current implementation.

The test statistic $t\left(\hat{\rho}_t^*, \kappa\right)$ has a $N(0,1)$ limiting distribution under the null hypothesis with a rejection region in the left tail.

5.A.3 Auxiliary results

Table 5.8: Hanck's (2013) and Demetrescu et al.'s (2006) panel unit root tests

Country	lag	ADF_τ	p-value	Country	lag	ADF_τ	p-value	Simes' crit. value	
	p				e			5%	10%
Italy	2	−3.97	0.014	Poland	2	−3.13	0.029	0.003	0.005
Austria	2	−3.66	0.031	Sweden	1	−2.92	0.047	0.005	0.011
Belgium	2	−3.41	0.057	Denmark	4	−2.01	0.282	0.008	0.016
Luxembourg	4	−3.11	0.110	Austria	1	−1.95	0.308	0.011	0.021
Denmark	5	−3.08	0.119	Lithuania	4	−1.86	0.347	0.013	0.026
Portugal	2	−2.75	0.221	Germany	1	−1.86	0.347	0.016	0.032
Poland	3	−2.72	0.234	Italy	1	−1.84	0.357	0.018	0.037
France	2	−2.67	0.250	France	1	−1.71	0.425	0.021	0.042
United Kingdom	1	−2.60	0.282	Belgium	1	−1.68	0.436	0.024	0.047
Sweden	1	−2.56	0.298	Finland	1	−1.68	0.440	0.026	0.053
Finland	5	−2.27	0.443	Netherlands	1	−1.67	0.442	0.029	0.058
Germany	2	−2.13	0.521	Czech Republic	2	−1.66	0.446	0.032	0.063
Estonia	6	−2.12	0.525	Luxembourg	1	−1.66	0.449	0.034	0.068
Lithuania	3	−2.04	0.572	Spain	1	−1.62	0.466	0.037	0.074
Netherlands	3	−1.86	0.664	Portugal	1	−1.61	0.473	0.039	0.079
Czech Republic	2	−1.84	0.673	Norway	1	−1.38	0.588	0.042	0.084
Norway	1	−1.84	0.677	United Kingdom	1	−1.22	0.661	0.045	0.089
Switzerland	4	−1.63	0.771	Switzerland	1	−0.65	0.853	0.047	0.095
Spain	2	−1.48	0.827	Estonia	1	−0.23	0.929	0.050	0.100
	y				w			5%	10%
Sweden	3	−3.44	0.054	Italy	2	−3.91	0.016	0.003	0.005
Italy	1	−3.03	0.132	France	1	−3.24	0.084	0.005	0.011
France	2	−3.03	0.132	Norway	3	−3.15	0.102	0.008	0.016
Germany	1	−3.02	0.134	Spain	2	−2.92	0.162	0.011	0.021
Luxembourg	2	−2.94	0.157	Germany	2	−2.75	0.220	0.013	0.026
Estonia	3	−2.92	0.161	Switzerland	2	−2.73	0.230	0.016	0.032
United Kingdom	1	−2.89	0.170	Netherlands	3	−2.72	0.231	0.018	0.037
Switzerland	1	−2.89	0.171	Czech Republic	4	−2.70	0.240	0.021	0.042
Belgium	1	−2.73	0.226	Estonia	4	−2.63	0.267	0.024	0.047
Austria	1	−2.64	0.265	Portugal	2	−2.62	0.274	0.026	0.053
Finland	3	−2.57	0.294	Finland	2	−2.57	0.296	0.029	0.058
Lithuania	3	−2.54	0.309	Austria	4	−2.35	0.401	0.032	0.063
Spain	6	−2.49	0.330	Poland	1	−2.33	0.411	0.034	0.068
Netherlands	2	−2.40	0.378	Belgium	1	−2.29	0.433	0.037	0.074
Denmark	2	−2.39	0.381	Lithuania	3	−2.11	0.530	0.039	0.079
Poland	4	−2.27	0.446	Luxembourg	2	−2.02	0.580	0.042	0.084
Portugal	1	−2.14	0.513	United Kingdom	5	−1.94	0.624	0.045	0.089
Czech Republic	1	−2.06	0.561	Denmark	2	−1.89	0.650	0.047	0.095
Norway	5	−2.04	0.571	Sweden	1	−1.12	0.919	0.050	0.100

Variable	Hartung's κ_2 test statistic
p	-0.804
e	-0.279
y	-0.719
w	-0.668

Notes: ADF_τ denotes the Augmented Dickey-Fuller test statistic. A linear time trend is included in the test regressions for p, y and w, while only a constant is considered for e. The lag order is selected according to the modified AIC (MAIC) criterion of Ng and Perron (2001). The p-values are computed as in MacKinnon (1996); the author is grateful to Christoph Hanck for the GAUSS code. Results for each variable are sorted according to the p-values in ascending order for ease of comparison with the corresponding Simes' critical value.

Table 5.9: Pesaran's (2007) $CIPS$ panel unit root test

lag	p	e	y	w	Δp	Δe	Δy	Δw
6	−2.037	−1.662	−2.322	−1.548	−3.216***	−3.708***	−2.550***	−2.659***
5	−2.022	−1.734	−2.210	−1.503	−3.378***	−4.134***	−2.807***	−3.071***
4	−2.049	−2.221**	−2.061	−1.571	−3.971***	−4.442***	−3.084***	−3.751***
3	−2.218	−2.460***	−2.267	−1.851	−4.315***	−3.880***	−3.768***	−4.300***
2	−2.517	−2.305**	−2.116	−2.050	−5.283***	−4.236***	−4.316***	−4.520***
1	−2.702**	−2.278**	−2.213	−1.857	−5.798***	−5.200***	−5.168***	−5.196***

Notes: Trend is included in the test regressions for p, y and w, while only a constant is considered for e. The 10%, 5% and 1% critical values for the model with constant only are −2.11, −2.2 and −2.36, and −2.63, −2.7 and −2.85 for the model with trend, respectively.
*, ** and *** denote significance at the 10, 5 and 1% level, respectively.

Table 5.10: Estimated factor loadings

Country	$\hat{\Lambda}_1^p$	$\hat{\Lambda}_2^p$	$\hat{\Lambda}_1^e$	$\hat{\Lambda}_2^e$	$\hat{\Lambda}^y$	$\hat{\Lambda}^w$
Austria	1.19	0.57	−1.26	−0.13	1.14	1.76
Belgium	1.25	−0.20	−1.27	−0.15	1.13	1.47
Czech Republic	1.06	0.23	−0.46	−1.71	1.15	0.17
Denmark	0.71	−1.87	−1.23	0.49	0.81	1.20
Estonia	0.94	−0.90	−0.83	1.52	0.95	0.84
Finland	1.08	−0.18	−1.25	0.25	1.11	−0.32
France	1.37	−0.08	−1.27	−0.28	1.22	1.36
Germany	1.35	−0.26	−1.27	−0.17	1.11	0.35
Italy	1.32	−0.04	−1.27	−0.20	1.23	0.57
Lithuania	0.56	1.78	−0.15	1.32	0.93	−0.06
Luxembourg	0.09	−2.32	−1.26	0.17	0.64	0.84
Netherlands	0.88	0.72	−1.27	−0.01	1.16	1.12
Norway	0.80	−0.14	−0.22	−1.52	0.38	1.11
Poland	0.30	1.60	−0.09	−2.05	0.36	0.38
Portugal	1.20	−0.79	−1.27	−0.06	0.94	1.09
Spain	1.24	−0.41	−1.27	−0.07	1.06	0.96
Sweden	0.98	0.72	−0.25	−1.90	1.03	−0.72
Switzerland	0.62	0.70	0.22	0.50	1.06	1.68
United Kingdom	0.86	0.84	0.60	−1.07	1.01	0.63

Figure 5.2: Estimated common factors from the Cup-FM model

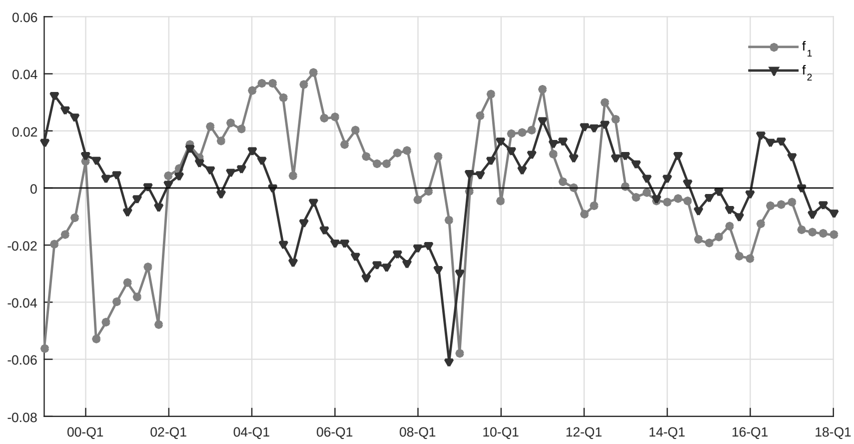


Figure 5.3: Estimated residuals from the Cup-FM model

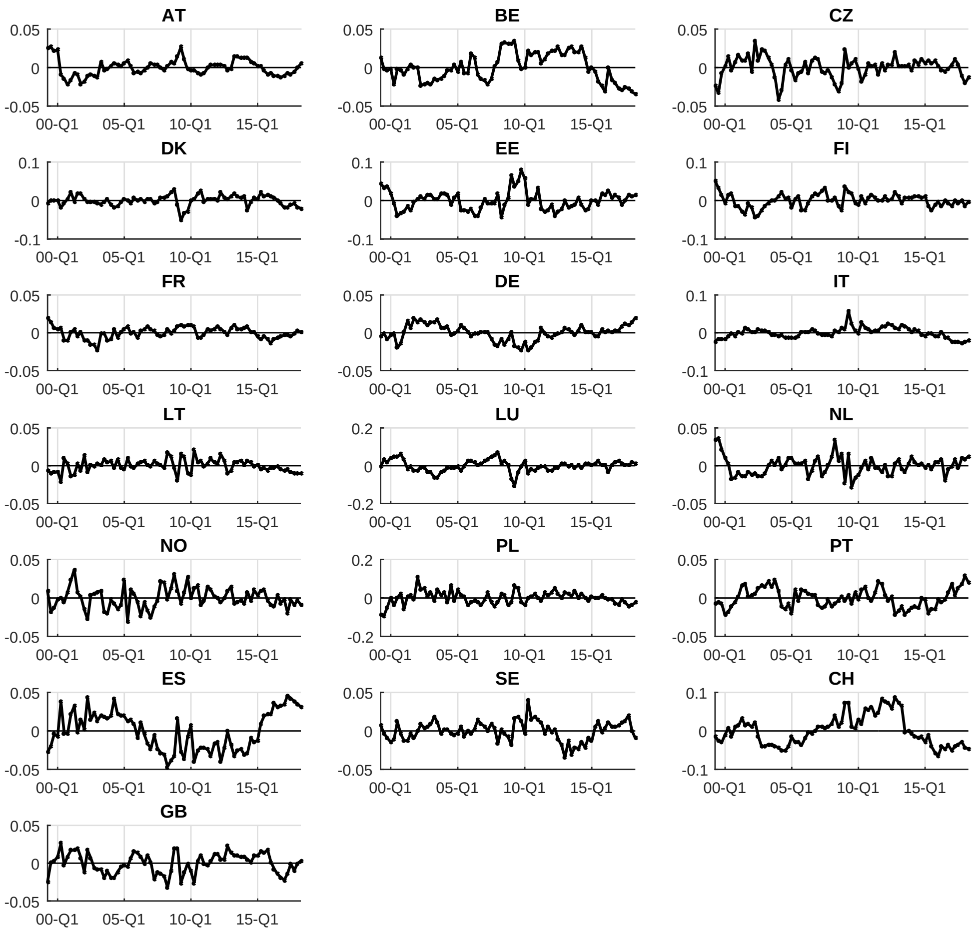


Figure 5.4: Estimated residuals from the augmented DCCE model with trend

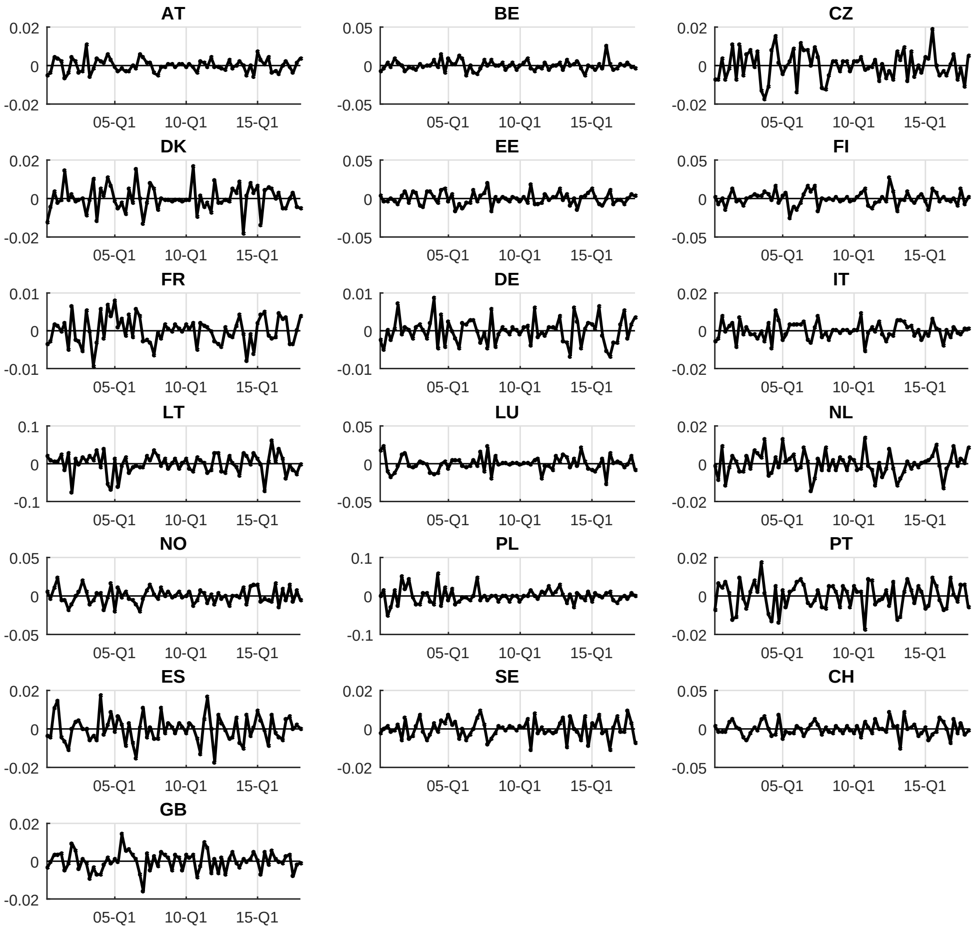


Table 5.11: ERPT estimation results for the individual countries by the DCCE estimator

	Exchange rate e		Domestic demand y		Producers' costs w		EC coefficient ρ_i
Country	long-run	short-run	long-run	short-run	long-run	short-run	
Austria	−0.11	−0.20	−0.15	0.34	0.01	0.06	−0.68***
	(0.17)	(0.20)	(0.25)	(0.23)	(0.12)	(0.13)	(0.19)
Belgium	0.08	−0.19	1.27	0.12	0.32	−0.04	−0.47***
	(0.49)	(0.37)	(0.89)	(0.63)	(0.38)	(0.27)	(0.15)
Czech Republic	−0.37***	−0.35***	−0.34	0.79	−0.46*	0.00	−0.83***
	(0.14)	(0.11)	(0.34)	(0.52)	(0.24)	(0.19)	(0.18)
Denmark	−1.46***	−1.11***	−0.42	0.30	−0.50***	0.01	−1.04***
	(0.39)	(0.30)	(0.30)	(0.34)	(0.15)	(0.21)	(0.17)
Estonia	−0.05	0.26	0.43**	0.32*	0.09	0.11	−0.88***
	(0.09)	(0.18)	(0.20)	(0.19)	(0.12)	(0.20)	(0.15)
Finland	−0.12	0.30	0.06	0.47	0.33	0.41	−0.80***
	(0.27)	(0.38)	(0.35)	(0.45)	(0.31)	(0.19)	(0.15)
France	0.03	0.29	0.40	0.00	0.21	0.05	−0.81***
	(0.14)	(0.21)	(0.47)	(0.42)	(0.16)	(0.17)	(0.18)
Germany	−0.51**	−0.44***	−0.33*	−0.06	0.10	0.22*	−0.64***
	(0.24)	(0.14)	(0.20)	(0.17)	(0.13)	(0.12)	(0.15)
Italy	−1.08**	−0.72***	1.00**	0.02	0.20	−0.04	−0.61***
	(0.45)	(0.23)	(0.44)	(0.35)	(0.28)	(0.14)	(0.16)
Lithuania	−1.17***	−0.98***	−0.62	−1.71**	0.52	−0.49	−0.82***
	(0.37)	(0.33)	(0.71)	(0.73)	(0.36)	(0.39)	(0.15)
Luxembourg	−0.16	−0.26	0.80	0.67*	−0.03	0.49*	−0.61***
	(0.67)	(0.68)	(0.55)	(0.35)	(0.26)	(0.26)	(0.13)
Netherlands	−0.03	−0.18	−0.53	−0.27	0.07	0.01	−0.87***
	(0.21)	(0.29)	(0.40)	(0.43)	(0.12)	(0.15)	(0.16)
Norway	−0.46*	−0.22**	−0.62	−0.18	−0.21	0.19	−0.57***
	(0.24)	(0.11)	(0.82)	(0.40)	(0.35)	(0.30)	(0.19)
Poland	−0.30**	−0.25	−0.08	−0.27	−0.75*	−0.53	−1.13***
	(0.13)	(0.19)	(0.31)	(0.62)	(0.42)	(0.34)	(0.20)
Portugal	−1.26**	0.20	1.18***	0.42	0.46*	0.26*	−0.71***
	(0.55)	(0.61)	(0.36)	(0.35)	(0.24)	(0.15)	(0.14)
Spain	0.64	0.48	0.63**	2.31***	0.41	−0.15	−0.84***
	(0.43)	(0.54)	(0.27)	(0.67)	(0.37)	(0.21)	(0.15)
Sweden	−0.82***	−0.41***	1.10***	−0.15	0.00	−0.16*	−0.80***
	(0.17)	(0.07)	(0.40)	(0.28)	(0.06)	(0.09)	(0.12)
Switzerland	0.65	−0.11	−2.61	1.38*	−12.25	1.04	−0.10
	(1.13)	(0.12)	(4.30)	(0.82)	(14.60)	(0.92)	(0.10)
United Kingdom	−0.47***	−0.31***	−0.38	−0.45	−0.32	−0.23*	−0.55***
	(0.18)	(0.06)	(0.36)	(0.36)	(0.24)	(0.12)	(0.15)

Notes: Results from the DCCE estimator with trend, augmented with a Great Recession dummy variable. Standard errors are presented in brackets.
*, ** and *** denote significance at the 10, 5 and 1% level, respectively.