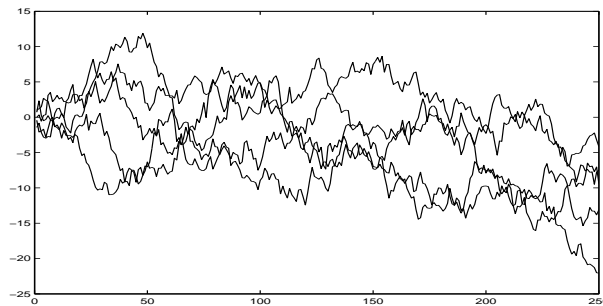


Unit Root Testing in Panel and Time Series Models: New Tests and Economic Applications



Inaugural-Dissertation
zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Dipl. Volkswirt Florian Siedenburg
aus Oldenburg

Kiel, 2010

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekanin:
Prof. Dr. Birgit Friedl

Erstberichterstattender:
Prof. Dr. Helmut Herwartz

Zweitberichterstattender:
Prof. Dr. Johannes Bröcker

Tag der Abgabe der Arbeit:
27. Januar 2010

Tag der mündlichen Prüfung:
16. Juli 2010

Für meine Familie: Jana, Franka und Luk.

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Statistik und Ökonometrie der Christian-Albrechts-Universität zu Kiel.

Mein besonderer Dank gilt meinem Doktorvater Professor Dr. Helmut Herwartz. Durch seine zahlreichen Anregungen, immerwährende Diskussionsbereitschaft und einen nie nachlassenden Enthusiasmus für unsere gemeinsamen Projekte trug er entscheidend zum Gelingen dieser Arbeit bei. Herrn Professor Dr. Johannes Bröcker danke ich herzlich für die Übernahme des Zweitgutachtens. Zu Dank verpflichtet bin ich außerdem Diplom-Informatiker Albrecht Mengel, der sämtliche vorstellbaren Computerprobleme lösen konnte.

Dankbar bin ich weiterhin allen Kollegen am Institut für Statistik und Ökonometrie. Sie haben nicht nur durch eine Vielzahl hilfreicher Anregungen im Rahmen des statistisch-ökonomischen Seminars zum Gelingen dieser Arbeit beigetragen, sondern auch durch eine angenehme Zusammenarbeit für ein produktives Arbeitsklima gesorgt. Hierbei ist insbesondere mein Freund, Diplom-Volkswirt Jan Roestel herauszuheben, der mich während der gesamten Zeit in unserem gemeinsamen Büro ertragen musste.

Schließlich möchte ich mich bei meiner Frau Jana dafür bedanken, dass sie mich durch die vielen Höhen und Tiefen der vergangenen vier Jahre liebevoll und verständig begleitet hat.

Kiel, am 6. Oktober 2010

Florian Siedenburger

Contents

1	Introduction	11
1.1	Outline	15
2	A Review of Unit Root Tests	19
2.1	Introduction	19
2.2	Preliminaries	20
2.3	Tests based on the first-order autoregression	21
2.3.1	The Dickey-Fuller test	21
2.3.2	The augmented Dickey-Fuller test	23
2.4	Semi- and nonparametric tests	24
2.4.1	The class of Z -tests	24
2.4.2	The class of M -tests	25
2.4.3	Fully nonparametric tests	27
2.4.3.1	A unit root test based on superfluous regressors	27
2.5	Other approaches to unit root testing	28
2.5.1	Tests based on quasi-differencing	28
2.5.2	Bootstrap unit root tests	29
2.5.2.1	Wild bootstrap unit root tests	30
2.6	Discussion of finite sample performances	31
2.7	Conclusions	32
3	A New Approach To Unit Root Testing	34
3.1	Introduction	34
3.2	The simulation based unit root test	36
3.2.1	The testing principle	36
3.2.2	Construction of the test statistic	37

CONTENTS

3.2.3	Consistent estimation of the long run variance	39
3.2.3.1	Lag length selection	40
3.2.4	Deterministic terms	41
3.2.5	Critical values	41
3.3	Finite sample properties	42
3.3.1	Rejection frequencies in the unit root case	43
3.3.2	Size adjusted local power	48
3.4	Empirical illustration: PPP of G7 economies	51
3.4.1	Theoretical background	51
3.4.2	Data	52
3.4.3	Results	53
3.5	Conclusions	54
4	A Review of Homogenous Panel Unit Root Tests	57
4.1	Introduction	57
4.2	The autoregressive panel model	58
4.3	First generation homogenous PURTs	60
4.3.1	The basic model	60
4.3.2	Deterministic Terms	61
4.3.3	Higher order serial correlation	63
4.4	Second generation homogenous PURTs	64
4.4.1	A feasible GLS-PURT	65
4.4.2	Robust covariance estimation	66
4.4.3	Bootstrap PURTs	67
4.5	Conclusions	68
5	Panel Unit Root Tests under Cross Sectional Dependence	69
5.1	Introduction	69
5.2	Cross sectional dependence in panel data	71
5.2.1	Unit root testing in the AR(1) panel model	71
5.2.2	Common factors	72
5.2.3	Spatial dependence	73
5.3	Finite sample modifications of homogenous PURTs	73
5.3.1	A ‘White’ correction	74
5.3.2	Improved finite sample residuals	75

CONTENTS

5.4	Monte Carlo study	76
5.4.1	The simulation design	76
5.4.2	Rejection frequencies under H_0	77
5.4.3	Rejection frequencies under H_1	82
5.4.4	Sensitivity analysis	82
5.5	Bootstrap inference	83
5.5.1	Setup of bootstrap algorithm	83
5.5.2	Asymptotic validity	84
5.5.3	Finite sample performance	86
5.5.4	Summary of Monte Carlo results	88
5.6	Empirical application	88
5.7	Conclusions	91
5.8	Appendix	93
5.8.1	Proof of Lemma 5.1	93
5.8.2	Proof of Proposition 5.3	94
5.8.3	Invariance principle for the wild bootstrap based on estimated residuals	95
6	Panel Unit Root Tests under a break in the innovation variance	96
6.1	Introduction	96
6.2	Effects of nonstationary volatility on univariate unit root tests	98
6.3	PURTs under nonstationary volatility	99
6.3.1	The autoregressive, heteroskedastic panel model	99
6.3.2	Asymptotic size distortions of homogenous PURTs	101
6.3.3	A volatility-break robust test	103
6.3.4	Local asymptotic power of t_{Wh}	104
6.4	Monte Carlo study	105
6.4.1	The simulation design	105
6.4.2	Results	107
6.4.3	Summary of simulation results	113
6.5	Testing the Fisher hypothesis by means of PURTs	114
6.5.1	Economic background	114
6.5.2	Data and preliminary analyses	115
6.5.3	Panel unit root test results	118
6.6	Conclusions	120

CONTENTS

6.7	Appendix	122
6.7.1	Proof of Proposition 6.1	122
6.7.2	Proof of Proposition 6.2	124
6.7.3	Proof of Proposition 6.3	125
6.7.4	Proof of Proposition 6.4	126
7	Testing convergence of unit labor costs in the Euro area: non-stationary volatility and wild bootstrap panel unit root tests	128
7.1	Introduction	128
7.2	PURTs under nonstationary volatility	130
7.2.1	The autoregressive, heteroskedastic panel model	130
7.3	Wild bootstrap PURTs	131
7.3.1	The bootstrap algorithm	132
7.3.2	Asymptotic properties of wild bootstrap PURTs	133
7.3.2.1	Asymptotic validity	133
7.3.2.2	Consistency	135
7.4	Monte Carlo study	136
7.4.1	The simulation design	136
7.4.2	Results	137
7.5	Persistent inflation differentials and stability of the EMU	142
7.5.1	Background	142
7.5.2	Data set and methodology	144
7.5.3	Results	149
7.6	Conclusions	155
8	Conclusions	158
	References	162

List of Figures

3.1	Empirical PDFs conditional on $R, T = 1000$	39
3.2	Real exchange rates of G7 economies	53
6.1	Asymptotic variance of t_{OLS}, \bar{v}_{OLS}	102
6.2	Nominal Interest Rates and Inflation rates, 1961Q2 - 2007Q2 . . .	116
6.3	Real interest rates, levels and 1st differences, US vs. UK	117
6.4	Estimated variance profiles	118
7.1	ULC growth differentials vis-a-vis the EU8 average	146
7.2	Volatility of the data	147
7.3	Country specific panels of ULC growth differentials	148
7.4	Country specific ULC growth differentials	150

List of Tables

3.1	Critical values, $J_{0.1}$	42
3.2	Empirical size	45
3.3	Empirical size, continued	47
3.4	Size adjusted local power	49
3.5	Size adjusted local power, continued	51
3.6	Empirical results	55
5.1	Homogenous PURTs under cross sectional independence	79
5.2	Homogenous PURTs under cross sectional correlation	81
5.3	Empirical results	90
6.1	DGP1	108
6.2	Empirical rejection frequencies, DGP2	111
6.3	Empirical rejection frequencies, DGP3	112
6.4	Interest rates, definitions	115
6.5	Empirical results	119
7.1	Finite sample properties: Rejection frequencies under H_0	138
7.2	Finite sample properties: Rejection frequencies under H_1	140
7.3	Descriptive statistics	149
7.4	Convergence of ULC growth differentials: PURT evidence	152
7.5	Convergence of ULC inflation differentials, univariate tests	154

Chapter 1

Introduction

Testing for unit roots is a fundamental building block of time series and panel data econometrics. The reason for this is twofold. First, it was demonstrated by Granger and Newbold (1974) that nonsensical correlations may arise from regressions involving two uncorrelated variables, integrated of order one, denoted as $I(1)$. Granger and Newbold (1974) consider the following stylized regression model

$$y_t = \beta_1 + \beta_2 x_t + u_t, \quad t = 1, \dots, T, \quad (1.1)$$

in which y_t and x_t are generated independently of each other as

$$y_t = y_{t-1} + \nu_t, \quad x_t = x_{t-1} + w_t, \quad \nu_t \sim iid(0, \sigma_\nu^2), \quad w_t \sim iid(0, \sigma_w^2).$$

Since y_t and x_t have autoregressive roots equal to one, they are commonly referred to as *unit root* processes. Granger and Newbold (1974) show that regressions such as (1.1) often lead to deceptive inference and misleading results. In particular, it is common to obtain seemingly highly significant estimates $\widehat{\beta}_2$ in combination with very high R^2 measures. However, these results come in conjunction with low Durbin-Watson statistics, indicating serially correlated regression residuals. Following Granger and Newbold (1974), regressions such as in (1.1) have been commonly labeled as *spurious* regressions. In order to avoid the nonsensical results obtained from spurious regressions, the econometric practitioner has to test for unit roots in the time series y_t and x_t in order to achieve a correctly specified regression model. If the time series turn out to be integrated of order one and there exists a long run (or so-called cointegration) relationship among the two variables, the model should be estimated in error correction form to make use

of the levels information of the data. If the series are $I(1)$ but uncorrelated in the long run, a regression in first differences avoids the spurious regression problem. A regression in levels is the appropriate model formulation if both variables turn out to be stationary or $I(0)$.

The second reason for the importance of unit root testing stems from economic theories, making use of equilibrium concepts. Economic equilibria are usually characterized by stable equilibrium values or growth paths. Only small deviations from these equilibrium values are permitted which are supposed to be promptly corrected by market forces. Theories making use of such equilibrium concepts can be tested by means of unit root tests: while unit root processes have no tendency to exhibit a stable mean value but rather follow some unpredictable stochastic path, stationary variables are characterized by a stable mean (or deterministic trend) around which the observations are fluctuating. Thus, unit root tests may be applied to directly test many (macro-) economic hypotheses. Prominent examples include among others the purchasing power parity (PPP) hypothesis, implying stationary real exchange rates, the Fisher hypothesis, postulating stationarity of real interest rates or growth convergence hypotheses. It has to be noted that fundamentally different policy recommendations may be obtained, depending on the order of integration of a specific variable. For instance, the question if GDP is an integrated or rather trend stationary variable has far reaching implications for the effects of discretionary policy actions.

Given the paramount importance of testing for unit roots in time series econometrics, it is not surprising that there is a huge literature on the topic. Starting with the early contributions of Dickey and Fuller (1979, 1981) who concentrate on testing the unit root hypothesis for random walks with uncorrelated increments, a large part of the literature on unit root testing deals with the treatment of serially correlated innovations. For instance, Said and Dickey (1984) propose to augment the Dickey-Fuller procedure by lags of the first difference of the data, where the lag length is chosen upon a data driven criterion. In contrast, semi-parametric approaches characterize the so-called Z - and M -type unit root tests (e.g. Phillips and Perron, 1988 or Stock, 1999). These tests account for serially dependent innovations by application of serial correlation robust nonparametric variance estimators. Fully nonparametric approaches (see, for instance, Breitung, 2002) as well as simulation and resampling based unit root tests (e.g. Cavaliere

and Taylor, 2007b, 2008) have been proposed in order to account for outliers, non-linear data transformations and breaks in the processes' unconditional volatility. While nonparametric approaches are usually more robust with respect to violations of assumptions underlying parametric tests, they are generally less powerful than parametric tests in finite samples.

Starting in the early 1990ies with the work of Levin et al. (2002)¹, Quah (1994) or Breitung and Meyer (1994), research on panel unit root tests (PURT) has seen a dramatic development in the recent past. PURTs have been proposed to overcome the evident power deficiency of univariate unit root tests. Especially for rather short time series, as often encountered in macroeconometrics with data typically sampled at the annual or quarterly frequency, univariate tests are found to suffer from low power against highly persistent stationary processes. These early PURTs are constructed as straightforward extensions of the univariate case by using the t -ratio obtained from a pooled Dickey-Fuller regression. Apart from yielding more powerful tests, the resulting test statistics follow a Gaussian limiting distribution. Such pooled PURTs are constructed against a homogenous alternative hypothesis under which equality of the autoregressive coefficient across panel units is imposed. However, since this assumption very unlikely holds in macroeconomic applications of interest, various authors (for instance, Maddala and Wu, 1999, Choi, 2001 or Im et al., 2003) have subsequently pursued the development of heterogenous PURTs. Heterogenous PURTs are formulated against a less restrictive alternative hypothesis where the autoregressive parameter may display heterogeneity across panel units, provided that a non-zero fraction of the cross sectional units displays mean reverting behavior. These heterogenous tests are either constructed by averaging over individual specific unit root tests statistics or by combining p -values obtained from univariate tests in the spirit of Fisher (1954).

The decision whether to use homogenous or heterogenous tests marks a significant turning point in the PURT literature. Im et al. (2003) and Maddala and Wu (1999) argue very strongly against the use of homogenous tests. While Maddala and Wu (1999) insist that the homogenous alternative hypothesis “[...] is too strong to be held in any interesting empirical cases” and hence argue in favor of heterogenous alternatives, Im et al. (2003) use statistical arguments to

¹The working paper version of this paper dates back to 1992.

dismiss the homogenous alternative. They justify their choice of the alternative hypothesis by recalling from Pesaran and Smith (1995) and Pesaran et al. (1995) that pooled estimators are inconsistent in dynamic heterogenous panel models. However, it has to be noted that inconsistency of the pooled autoregressive parameter estimate not necessarily affects the consistency of homogenous PURTs (Hurlin and Mignon, 2007). In fact, it has been pointed out by Moon and Perron (2004), Moon and Perron (2008) or Breitung and Pesaran (2008) that pooled tests are consistent against heterogenous alternatives and are indeed even more powerful than their heterogenous counterparts. Yet, among practitioners there appears to be still some degree of confusion when it comes to interpreting rejections of the panel unit root hypothesis. While some authors tend to (falsely) interpret a rejection of heterogenous tests as indicative of a fully stationary panel, other authors apply homogenous tests by arguing that the homogenous alternative is economically more meaningful². More recently though, authors who apply PURTs in empirical studies seem to be increasingly aware that the correct interpretation of a rejection of the panel unit root hypothesis is to “conclude that a significant fraction of the AR(1) processes in the panel does not contain unit roots” (Breitung and Pesaran, 2008).

While the distinction between homogenous and heterogenous tests can be seen as a first division of the PURT literature, a more significant innovation can be identified with respect to the assumptions regarding contemporaneous correlation among the panel units. Early (the so-called first generation) PURTs are constructed around the assumption that cross sectional units are contemporaneously uncorrelated. This assumption simplifies the analytical derivations considerably by allowing the application of standard central limit theorems for uncorrelated random variables. However, since cross sectional correlation is a typical characteristic of most macroeconomic data sets, this assumption is very restrictive and it turned out that first generation PURTs effectively lose size control if the data exhibits cross sectional correlation. Accordingly, a new (or second) generation of PURTs emerged, which explicitly allows for cross sectional dependence. This literature can also be roughly divided into two directions. First, there is a large number of authors who presume that the cross sectional dependence stems from the class of common factor models (Phillips and Sul, 2003, Bai and Ng,

²In the early stages of this work, the author has also been an adherent of this view.

2004, Moon and Perron, 2004, Choi, 2006, or Pesaran, 2007). The second group consists of tests which are constructed against unknown forms of cross sectional dependence by imposing only few restrictions on the covariance matrix of regression residuals. Examples of the latter group are given by, for instance, Chang (2002), Harvey and Bates (2003), Chang (2004), Jönsson (2005) or Breitung and Das (2005).

1.1 Outline

It is the goal of this thesis to contribute to the literature of testing for unit roots, both in time series as well as in panel data models. The main part of the thesis consists of six chapters. While Chapters 2 and 3 deal with univariate unit root tests, the focus of Chapters 4 through 7 is on the fast growing research area of testing for unit roots in panel data models. Chapters 2 and 4 are literature surveys on univariate- and panel unit root tests, respectively, and Chapters 3, 5, 6 and 7 are adapted versions of original research papers coauthored by Helmut Herwartz. While the contribution of Chapter 3 is to provide a not yet explored route of unit root testing, the motivation of Chapters 5 through 7 is to improve upon the finite sample properties or robustness of more established tests. Hereby, the focus is on developing pooled test procedures which do not rely on specific (parametric) assumptions regarding cross sectional error correlation. Particular attention is placed on analyzing the effects of breaks in the unconditional variance of the data generating process. Whilst a growing number of authors (e.g. Hamori and Tokihisa, 1997, Kim et al., 2002, Cavaliere, 2004 or Cavaliere and Taylor, 2007b, 2008, 2009) has considered this for the univariate case, variance breaks have so far been neglected in the literature of testing for unit roots in panel data models. Accordingly, this work has a rather methodological emphasis. However, empirical applications complement the theoretical expositions to demonstrate that results of applied research may indeed be sensitive with respect to the chosen tests. In order to make the chapters self containing and enhance readability, considered models as well as underlying assumptions are introduced separately in each of the chapters. Similarly, abbreviations used are also introduced afresh.

Chapter 2 provides a review over the literature on univariate unit root testing. The focus of the chapter is on the most commonly applied unit root test

procedures, such as the (augmented) Dickey-Fuller test and the classes of Z - and M -type unit root tests. However, to provide the basis for later chapters, non-parametric and bootstrap unit root tests are also briefly covered. In particular, the nonparametric test discussed in Chapter 2.4.3.1 is closely related to the simulation based unit root test suggested in Chapter 3. Both tests utilize the spurious regression results by regressing the data on computer generated random walks. Similarly, the wild bootstrap approach as discussed in Chapter 2.5.2.1 is reconsidered for the construction of a robust second generation PURT in Chapters 5 and 7.

A new simulation based unit root test is proposed in Chapter 3. The suggested test is built upon the fact that the estimated regression coefficient from a spurious regression does not converge to the true value of zero but rather has a nondegenerated limiting distribution. Similarly as the nonparametric test proposed by Park (1990), this characteristic is exploited by regressing the data on a computer generated random walk. Unlike the test of Park (1990), which is constructed from a single regression, the test developed in Chapter 3 rests on a sequence of such regressions. Under the unit root null hypothesis, the resulting regression coefficients display some finite range while, under the alternative, this range becomes arbitrary small. The new test is shown to offer quite robust size control in a variety of model settings including serially correlated innovations of the autoregressive and moving average types and is competitive in terms of size adjusted local power. As an economic application, the PPP hypothesis in its weak form is reconsidered for a data set of the G7 economies. In line with much of the empirical literature on weak form PPP, the results in Chapter 3 provide evidence in favor of weak form PPP only for a minority of the sample economies.

To introduce the reader to the field of homogenous PURTs, Chapter 4 provides a review over the relevant literature. First, the autoregressive panel model is discussed. Subsequently, variants of first generation PURTs based on a pooled Dickey-Fuller regression are presented. These tests differ mostly with respect to the treatment of deterministic terms. While Levin et al. (2002) propose to remove the impact of deterministic terms by least squares demeaning or detrending, this approach involves complicated bias adjustment terms in the resulting test statistic. Breitung and Meyer (1994) and Breitung (2000) suggest tests which avoid these bias adjustment terms by applying alternative demeaning or detrending

schemes. The discussion of second generation PURTs focuses on such procedures which do not assume a particular model of cross sectional correlation. A feasible GLS-PURT suggested by Harvey and Bates (2003) offers preferable asymptotic properties but only works in particular settings where the time dimension is significantly larger than the cross sectional dimension. Subsequently, an alternative approach independently suggested by Jönsson (2005) and Breitung and Das (2005) is reviewed. This test is based on an OLS Dickey-Fuller regression and cross sectional dependence is accounted for by application of so-called panel corrected standard errors (Beck and Katz, 1995). Finally, some bootstrap panel unit root tests are considered.

In Chapter 5, a modified PURT statistic is suggested. As the statistic proposed by Breitung and Das (2005), it is also based on a pooled Dickey-Fuller regression. However, instead of panel corrected standard errors, a panel adaptation of the White-heteroskedasticity consistent standard errors (White, 1980) is applied in the construction of the statistic. The applied variance estimator is constructed under the null hypothesis, making use of the first differences of the data instead of regression residuals. For models with strong form cross sectional dependence or a finite cross section dimension, the wild bootstrap is proposed as a general means of obtaining correct inference. The asymptotic properties of the proposed statistics are formally derived and the finite sample properties are assessed via a Monte Carlo simulation study. As an empirical application, the unit root hypothesis is tested for the current account to GDP ratio for a data set featuring up to 129 economies.

Chapters 6 and 7 focus on PURTs under discrete breaks in the unconditional volatility of the considered processes. As mentioned above, a growing literature has documented the adverse effects induced by nonstationary volatility on the performance of univariate unit root tests. To the knowledge of the author, the analyses in Chapters 6 and 7 are the first attempts to extend this literature to the panel case. Since both methods proposed in Chapter 5 - the ‘White’ corrected PURT statistic as well as the wild bootstrap approach - originate from robustifying inference in classical regression models against heteroskedasticity of unknown form, the performance of these PURTs under volatility breaks is assessed. In Chapter 6, it is shown that sudden shifts of the unconditional volatility of the process lead to marked size distortions of the PURTs suggested by Levin et al.

(2002) and Breitung and Das (2005). In contrast, the test suggested in Chapter 5, based on the ‘White’ corrected test statistic is proved to be robust against the considered break scenarios. As an economic application, the Fisher hypothesis is revisited. It can be empirically assessed by testing for a unit root in the real interest rates, which are shown to exhibit time varying volatility.

Finally, Chapter 7 reconsiders the wild bootstrap PURTs proposed in Chapter 5. It is recommended as an approach to overcome problems associated with a small cross sectional dimension, strong contemporaneous correlation and nonstationary volatility. The theoretical results from Chapter 5 are extended to the case of nonstationary volatility. As an extension to Chapter 6, a more general class of nonstationary volatility is considered here. A Monte Carlo simulation study confirms the theoretical arguments and provides evidence of the high degree of size control offered by the wild bootstrap PURTs. The remainder of Chapter 7 is devoted to an empirical assessment of unit labor cost developments among member countries of the European Economic and Monetary Union (EMU). It is argued that divergent unit labor costs developments pose a long term threat to the stability of the monetary union. In the tradition of the growth convergence literature (e.g. Barro and Sala-i-Martin, 1991 or Bernard and Durlauf, 1996), convergence of unit labor cost inflation across the considered economies is measured in terms of absolute β -convergence. This implies that unit labor cost inflation differentials between member countries should follow mean-zero stationary processes. The data set used for the empirical analysis consists of monthly data for eight Euro member economies. The convergence analysis is separately conducted for a pre and a post Euro introduction sample. For both subsamples, it is shown that there is considerable variation in the volatility of the data. Combined evidence from panel and univariate unit root tests hints at widening competitive positions among EMU economies since the introduction of the Euro. In particular, Germany has successively been improving its competitive position, whereas the relative competitive positions of Italy and Spain have been deteriorating.

Chapter 2

A Review of Unit Root Tests

2.1 Introduction

Dealing with nonstationary behavior is a central issue in time series econometrics. In the following, the term *nonstationarity* is used to characterize stochastic processes which are not *covariance* stationary or *weakly* stationary. A time series is called (weakly) stationary, if neither its unconditional mean nor its autocovariance (at any horizon) is a function of time. According to this definition of stationarity, most macroeconomic and financial time series exhibit nonstationary behavior: many aggregated series as, for example, gross domestic product or industrial production tend to evolve around a secular growth trend. Asset prices and interest rates often appear to be randomly wandering around with no fixed population mean. Finally, while arguably fluctuating around a stable mean, asset returns regularly display time dependent volatility clusters.

Since the work of Granger and Newbold (1974), it is a well known fact that problems can arise in a classical regression framework with nonstationary variables which are integrated of order one (in the following written as $I(1)$). A time series is said to be integrated of order one if its first difference is weakly stationary. This is equivalent with stating that an $I(1)$ process has an autoregressive root equal to one, hence such variables are commonly referred to as *unit root* processes. It follows accordingly that every weakly stationary process is integrated of order zero (or $I(0)$) but not every nonstationary variable has to be generated by a unit root process. Among the most important findings of Granger and Newbold (1974) is the observations that conventional significance tests based on

ordinary least squares (OLS) regressions of two uncorrelated $I(1)$ variables tend to spuriously indicate a high degree of correlation among two completely uncorrelated variables. Granger and Newbold (1974) also point out that these so-called spurious regressions often lead to very high R^2 measures in combination with highly autocorrelated residuals as indicated by very low Durbin-Watson statistics. Hence, in order to avoid the pitfalls associated with unit root processes in the classical regression framework, testing for unit roots has become an important branch in the time series econometric literature. It is the purpose of this Chapter to give a concise review over this branch of theoretical research. Due to the immense scope of the literature, the focus of this Chapter is narrow, concentrating on the most commonly applied univariate unit root tests as well as on approaches to unit root testing which are reconsidered in later Chapters of this thesis.

2.2 Preliminaries

Assume a time series y_t can be decomposed into a deterministic component d_t and a stochastic component x_t , such that

$$y_t = d_t + x_t, \quad x_t = \rho Lx_t + u_t, \quad t = 1, \dots, T, \quad (2.1)$$

where $d_t = \mu'z_t$ and z_t is a $p \times 1$ vector of deterministic components and μ is the corresponding $(p \times 1)$ parameter vector. The stochastic component x_t is assumed to be an autoregressive process of order one (AR(1)) with parameter ρ and L is the lag operator, such that $Lx_t = x_{t-1}$. It is assumed that u_t is a mean-zero disturbance term with finite variance, σ_u^2 . Further assumptions on the error term and stability conditions are often test specific and are thus stated in conjunction with the respective statistics. Moreover, it is usually assumed that the initial value x_0 is either known or drawn from some stationary probability distribution. The null hypothesis of all tests considered in the following is that y_t is an $I(1)$ process, which is equivalent to stating that $|\rho| = 1$. Since negative unit root processes are rarely encountered in economic applications, unit root tests are generally constructed against the positive part of the null hypothesis, $\rho = 1$, under which it holds that $\Delta x_t = u_t$ where Δ is the difference operator, $\Delta x_t = x_t(1 - L)$. The one sided alternative hypothesis of all tests is that y_t is

2.3 Tests based on the first-order autoregression

stationary or, alternatively, $\rho < 1$. The explosive case of $\rho > 1$ is ruled out by assumption.

An important quantity for the derivation of the test statistics can be constructed from the sequence of partial sums $\{x_t\}_{t=1}^T$ as

$$R_T(r) = T^{-1/2} \sigma_u^{-1} x_{[Tr]} = T^{-1/2} \sum_{t=1}^{[Tr]} u_t / \sigma_u, \quad (2.2)$$

where $r \in [0, 1]$ and $[Tr]$ denotes the integer part of Tr . The quantity $R_T(r)$ is a real valued random variable with support $[0, 1]$. Donsker (1951) states a *functional central limit theorem* (FCLT) or *invariance principle* for the asymptotic behavior of the random variable $R_T(r)$. If the error terms $\{u_t\}_{t=1}^T$ are independently and identically distributed (*iid*), then for $T \rightarrow \infty$,

$$R_T(r) \xrightarrow{d} W(r), \quad (2.3)$$

where $W(r)$ is a standard Brownian motion (see Billingsley, 1968 for a proof and extensions) and \xrightarrow{d} denotes *weak* convergence (or, alternatively, convergence *in distribution*). To simplify notation, the following shorthand notations are introduced which will be used throughout. W abbreviates $W(r)$ and $W_1 = W(1)$. Moreover, the integral $\int_0^1 W(r) d(r)$ is written as $\int W$. If not otherwise stated, all integrals are taken over the interval $[0, 1]$.

2.3 Tests based on the first-order autoregression

2.3.1 The Dickey-Fuller test

Dickey and Fuller (1979, 1981) propose to test the unit root null hypothesis directly, based on the first order autoregression of y_t on y_{t-1} . There are three possible regression models of interest,

$$y_t = \rho y_{t-1} + u_t \quad (2.4)$$

$$y_t = \alpha + \rho y_{t-1} + u_t \quad (2.5)$$

$$y_t = \alpha + \delta t + \rho y_{t-1} + u_t. \quad (2.6)$$

2.3 Tests based on the first-order autoregression

In all three regressions the unit root null hypothesis is tested by means of the OLS estimate of the autoregressive parameter, $\hat{\rho}$. The difference between the three regressions is the presumed presence of deterministic terms under the alternative hypothesis. Under regression (2.4), it is assumed that the process is stationary around zero under the alternative. Likewise, stationarity around some non-zero mean α is the alternative hypothesis under regression (2.5). Finally, trend stationarity is the alternative hypothesis considered in case (2.6). Omitting a deterministic component in the test regression leads to invalid results while superfluous inclusion of deterministic regressors does not invalidate the inference on $\hat{\rho}$ but may adversely affect the power of the test.

The tests proposed by Dickey and Fuller (1979, 1981) rely on the following strong assumption of white noise error terms:

Assumption 2.1 (\mathcal{A}_1)

The error terms are a sequence of white noise, $u_t \sim iid(0, \sigma_u^2)$.

Assumption \mathcal{A}_1 is a strong assumption since it rules out any type of serial correlation, heteroskedasticity and restricts the errors to originate from the same probability distribution for all time points t . Dickey and Fuller (1979, 1981) consider two particular statistics to test the null hypothesis of $\rho = 1$. The first statistic is directly based on the autoregressive parameter estimate and given as

$$DF_{\hat{\rho}} = T(\hat{\rho} - 1), \quad (2.7)$$

while the second is given as the according t -ratio,

$$DF_t = \frac{(\hat{\rho} - 1)}{\left(s_u^2 / \sum_{t=1}^T y_{t-1}^2\right)^{1/2}}, \quad (2.8)$$

where $s_u^2 = \sum_{t=1}^T \hat{u}_t^2 / (T - 1)$ is the OLS estimate of the residual variance σ_u^2 . The limiting distribution of $DF_{\hat{\rho}}$ and DF_t are both nonstandard, depend on the deterministic terms included in the test regression and can be expressed as functionals of Brownian motion. They are given as

$$DF_{\hat{\rho}} \xrightarrow{d} \left[\int W_Z dW \right] \left[\int W_Z^2 \right]^{-1}, \quad (2.9)$$

and

$$DF_t \xrightarrow{d} \left[\int W_Z dW \right] \left[\int W_Z^2 \right]^{-1/2}, \quad (2.10)$$

2.3 Tests based on the first-order autoregression

where $W_Z = W - \int W Z' (\int Z Z')^{-1} Z$ is the projection residual of W on the matrix of deterministic terms Z , in the Hilbert space $L_2[0, 1]$. In the most simple case of no deterministic terms in the test regression, these expressions simplify to

$$DF_{\hat{\rho}} \xrightarrow{d} \frac{1/2\{W_1^2 - 1\}}{\int W^2},$$

and

$$DF_t \xrightarrow{d} \frac{1/2\{W_1^2 - 1\}}{\{\int W^2\}^{1/2}},$$

respectively. Both distributions are asymmetric, with negative values twice as likely as positive values.

Phillips (1987) shows that if the tests based on regressions (2.4)-(2.6) are applied to series in which the innovation sequence $\{u_t\}_{t=1}^T$ is serially dependent, there are nuisance parameters in the limiting distributions. Thus, various tests have been proposed to account for serially correlated error terms.

2.3.2 The augmented Dickey-Fuller test

Dickey and Fuller (1981) show that the DF test remains valid for higher order autoregressive models, provided the correct autoregressive order is known, by augmenting the test regression by the appropriate number of lagged differences of y_t . As a generalization for the case where the innovations u_t follow an autoregressive-moving average (ARMA) processes of unknown order, Said and Dickey (1984) propose the augmented DF (ADF) procedure. Since, under H_0 , $\Delta y_t = u_t$, and abstracting from deterministic terms, they consider the modified test regression

$$\Delta y_t = \phi y_{t-1} + \sum_{j=1}^k \beta_j \Delta y_{t-j} + e_t, \quad t = k + 1, \dots, T, \quad (2.11)$$

where $\phi = \rho - 1$ and e_t is a white noise error terms. In regression (2.11), the unit root hypothesis $\phi = 0$ is tested against the alternative $\phi < 0$. The inclusion of the lagged differences of the dependent variable are supposed to approximate the residual serial correlation induced by ARMA models of unknown order if the chosen lag length k increases with the sample size T . In particular, Said and Dickey (1984) show that if k growth at a rate of less than $T^{1/3}$, the t -ratio of the

parameter estimate $\hat{\phi}$ has the same limiting distribution as the corresponding DF statistic (2.10).

2.4 Semi- and nonparametric tests

2.4.1 The class of Z -tests

A class of semiparametric unit root tests has been developed by Phillips (1987) and Phillips and Perron (1988). The tests are similar to the DF tests in that they are based on the first order autoregressions (2.4)-(2.6). However, unlike Said and Dickey (1984), nonparametric variance estimators are employed to account for general forms of serial correlation and heteroskedasticity. These tests are derived under the following weak assumptions on u_t .

Assumption 2.2 (\mathcal{A}_2)

- (i) $E(u_t) = 0$ for all t ;
- (ii) $\sup_t E|u_t|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$;
- (iii) as $T \rightarrow \infty$, $\sigma^2 = \lim E(T^{-1}S_T^2) > 0$, where $S_T = \sum_{t=1}^T u_t$;
- (iv) $\{u_t\}$ is strong mixing with mixing coefficients α_m that satisfy $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$.

The set of conditions summarized by Assumption 2 allows for a wide range of error processes which are extensively discussed in Phillips (1987). In the special case of *iid* errors, the so-called long run variance, $\sigma^2 = \sigma_u^2$ and the results of Dickey and Fuller (1979, 1981) apply.

Phillips (1987) and Phillips and Perron (1988) advocate modified variants of the statistics (2.7) and (2.8) which employ nonparametric long run variance estimators (Newey and West, 1987) to remove the nuisance parameters induced by serial correlation and/or heteroskedasticity. Define $\lambda = \frac{1}{2}(\sigma^2 - \sigma_u^2)$. Then, the modified test statistics based on regressions (2.4)-(2.6) are given by

$$Z_{\hat{\rho}} = T(\hat{\rho} - 1) - \hat{\lambda} \left(T^{-2} \sum_{t=2}^T y_{Z,t-1}^2 \right)^{-1} \xrightarrow{d} \left[\int W_Z dW \right] \left[\int W_Z^2 \right]^{-1}, \quad (2.12)$$

and

$$Z_t = (s_u/s)DF_t - \hat{\lambda} \left\{ s \left(T^{-2} \sum_{t=2}^T y_{Z,t-1}^2 \right)^{1/2} \right\} \xrightarrow{d} \left[\int W_Z dW \right] \left[\int W_Z^2 \right]^{-1/2}, \quad (2.13)$$

where s^2 is a consistent estimator of the long run variance σ^2 and $y_{Z,t}$ is the residual from a regression of y_t on the set of deterministic terms z_t . It is obvious from the limiting representations in (2.12) and (2.13) that the modified statistics share the same limiting distribution as the standard DF statistics, which allows to use the same tabulated critical values.

While there are various possible estimators for the long run variance σ^2 , Phillips and Perron (1988) advocate the use of the serial correlation and heteroskedasticity consistent variance estimator proposed by Newey and West (1987). It is given as

$$s_{NW}^2 = T^{-1} \sum_{t=1}^T \hat{u}_t^2 + 2T^{-1} \sum_{\tau=1}^k w_\tau \sum_{t=\tau+1}^T u_t u_{t-\tau}, \quad (2.14)$$

where the weight function w_τ is given by

$$w_\tau = 1 - \tau/(k + 1).$$

The specific choice of the Newey-West procedure can be motivated by noting that it ensures a nonnegative variance estimate. Moreover, for any weakly stationary series $\{u_t\}_{t=1}^T$, the estimator in (2.14) is simply 2π times the corresponding Bartlett estimate of the spectral density at frequency zero.

2.4.2 The class of M -tests

All tests presented so far directly examine the unit root hypothesis by means of the autoregressive parameter estimate $\hat{\rho}$, either directly or by its corresponding t -ratio. In contrast, the class of M -tests for integration and cointegration proposed by Stock (1999)¹ is based on the implication that an integrated process has a growing variance, that is, has a higher order in probability than a stationary process. However, Stock (1999) shows that tests based on the latter principle can

¹The paper dates back to 1990.

2.4 Semi- and nonparametric tests

be expressed as modified (hence the term M -tests) versions of many previously proposed tests. For example, consider the following statistics

$$MZ_{\hat{\rho}} = (T^{-1}y_T - s^2) \left(2T^{-2} \sum_{t=1}^T y_{t-1}^2 \right)^{-1} \quad (2.15)$$

and

$$MSB = \left(T^{-2} \sum_{t=1}^T y_{t-1}^2 / s^2 \right)^{1/2}. \quad (2.16)$$

It can be shown that the statistic (2.15) can be expressed as $MZ_{\hat{\rho}} = Z_{\hat{\rho}} + (T/2)(\hat{\rho} - 1)^2$, which is a modified version of the statistic $Z_{\hat{\rho}}$ given in (2.12), where $(T/2)(\hat{\rho} - 1)^2$ is the modification factor. As $\hat{\rho}$ converges to 1 at the rate of T under the null hypothesis, the critical values for $Z_{\hat{\rho}}$ apply for $MZ_{\hat{\rho}}$. Similarly, the statistic MSB can be seen as a modified version of the so-called R unit root test statistics proposed by Bhargava (1986) which is built upon the work of Sargan and Bhargava (1983). Critical values for the cases of demeaned and detrended y_t are tabulated in Stock (1999). Finally, Perron and Ng (1996) point out that $Z_t = MSB \cdot Z_{\hat{\rho}}$ and hence suggest to use the relationship $MZ_t = MSB \cdot MZ_{\hat{\rho}}$ to propose a modified version of the Z_t statistic (2.13)

$$MZ_t = Z_t + (1/2) \left(\sum_{t=1}^T y_{t-1}^2 / s^2 \right)^{1/2} (\hat{\rho} - 1)^2. \quad (2.17)$$

Regarding the estimation of the long run variance σ^2 , Perron and Ng (1998) propose an autoregressive spectral density estimator s_{AR}^2 as an alternative to the Newey-West estimator (2.14) employed by Phillips and Perron (1988). It is given as

$$s_{AR}^2 = s_{ek}^2 / (1 - \hat{\beta}(1))^2, \quad (2.18)$$

where $s_{ek}^2 = (T - k)^{-1} \sum_{t=k+1}^T \hat{e}_t^2$ and $\hat{\beta}(1) = \sum_{j=1}^k \hat{\beta}_j$ and \hat{e}_t and $\hat{\beta}_j$ are obtained from an ADF regression as given in (2.11). It is shown by means of a Monte Carlo study (Perron and Ng, 1996) that empirical rejection frequencies under the null hypothesis of Z - and M -type unit root tests are closer to the nominal level if the estimator s_{AR}^2 is used instead of s_{NW}^2 .

2.4.3 Fully nonparametric tests

Several authors have developed fully nonparametric tests which do not rely on a parametric regression model to capture deterministic terms or short run dynamics of the observed time series. Examples of this strand of the literature are, for instance, the tests suggested by Breitung and Gouriéroux (1997) or Aparico et al. (2006). Breitung and Gouriéroux (1997) provide test statistics which are computed on the ranks of the observations instead using the actual observations. Aparico et al. (2006) propose a range unit root test which is constructed as a scaled sum of the number of changes of the observed time series' range. The advantage of such nonparametric approaches is a less sensitive reaction against mis-specification of the assumed data generating process (DGP). For instance, outliers or nonlinear data transformations may lead to severe size distortions of parametric tests. On the hand, the power of parametric tests against stationary processes with breaks in the unconditional mean is usually significantly depressed. However, this advantage of nonparametric tests' increased robustness usually comes at the cost of reduced power if the DGP is adequately described by the parametric model. In the following, the nonparametric test of Park (1990) is described in detail as the exposition in Chapter 3 is closely related. For a more extensive discussion of nonparametric unit root tests, the reader is referred to Breitung (2002) who reviews several nonparametric tests for unit roots and cointegration.

2.4.3.1 A unit root test based on superfluous regressors

Park et al. (1988), Park and Choi (1988) and Park (1990) develop a nonparametric approach to unit root testing which builds upon the spurious regression results derived by Phillips (1986). The unit root test suggested in Park (1990) is based on the following test regression

$$y_t = \tilde{\beta}' s_t + e_t, \quad (2.19)$$

where $s_t = (s_{1t}, \dots, s_{mt})'$ is an $m \times 1$ computer generated Gaussian random walk and $\tilde{\beta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_m)'$ is the corresponding $(m \times 1)$ vector of OLS parameter estimates. If, under the unit root null hypothesis, y_t is $I(1)$, then the elements in $\tilde{\beta}$ do not converge in probability towards the true parameters $\beta = (0, \dots, 0)'$ but converge weakly to a nonstandard limiting distribution. A consistent test statistic

2.5 Other approaches to unit root testing

which is asymptotically free of nuisance parameters, and hence does not require any adjustment for short run dynamics, can be constructed as a standardized Wald statistic $F(\tilde{\beta})$ to test the hypothesis $H_0 : \tilde{\beta} = (0, \dots, 0)'$ in regression (2.19). Scaling $F(\tilde{\beta})$ by the number of observations yields the nonparametric test statistic

$$J_2(m) = T^{-1}F(\tilde{\beta}) = T^{-1} \frac{RSS_1 - RSS_2}{RSS_2}, \quad (2.20)$$

where $RSS_1 = \sum_{t=1}^T y_t^2$ and $RSS_2 = \sum_{t=1}^T (y_t - \tilde{\beta}'s_t)^2$ are the residual sums of squares of the constrained and unconstrained regression, respectively². It follows from the results in Park (1990) that under H_0 , the limiting distribution of $J_2(m)$ can be expressed as

$$J_2(m) \xrightarrow{d} \frac{\int W^2 - \int U^2}{\int U^2}, \quad (2.21)$$

where W is a standard Brownian motion and $U = W - \int WW'_m (\int W_m W'_m) W_m$ is the projection residual in $L_2[0, 1]$ of W on the m dimensional Brownian motion W_m associated with s_t . Consistency of the test follows from the fact that under the alternative hypothesis, $J_2(m) \xrightarrow{p} 0$ at the rate T . The test is one sided and critical value have to be obtained by simulation for any specific choice of the dimension m . As discussed in Park (1990) and Park and Choi (1988), the choice of the dimension m has some impact on the test's power in finite samples. Park (1990) proposes to use at least $m = 2$, while Breitung (2002) employs the J_2 statistic with $m = 4$. To account for a non-zero intercept or a deterministic time trend, a vector of deterministic terms is included in the test regression.

2.5 Other approaches to unit root testing

2.5.1 Tests based on quasi-differencing

Noting that the (A)DF tests are characterized by a significant loss of asymptotic power in the cases (2.5) and (2.6) compared with case (2.4), Elliott et al. (1996) propose a generalized least squares (GLS) detrending scheme as an alternative

²Park (1990) proposes two separate statistics, one for the null hypothesis of stationarity and the the other for the unit root null hypothesis. While the further is named J_1 , the latter is accordingly called J_2 .

2.5 Other approaches to unit root testing

which avoids this asymptotic power loss. In particular, they propose to run the following regression

$$y_t^{\bar{\rho}} = \mu^{\bar{\rho}'} z_t^{\bar{\rho}} + u_t^{\bar{\rho}}, \quad (2.22)$$

where $\bar{\rho} = 1 + \bar{c}/T$ is the local detrending parameter and $(y_0^{\bar{\rho}}, y_t^{\bar{\rho}}) = (y_0, (1 - \bar{\rho}L)y_t)$ for all $t = 1, \dots, T$ and $z_t^{\bar{\rho}}$ is constructed accordingly. For the choice of \bar{c} , Elliott et al. (1996) suggest the values -7 and -13.5 for the case of an intercept and a linear trend, respectively. These values are chosen such that the asymptotic power envelope is tangent to 50%. Note, however, that these values apply only if the initial value $y_0 = 0$. If the initial value is drawn from its unconditional distribution, Elliott (1999) shows that different values for \bar{c} have to be chosen to obtain tangency of the power envelope to the 50% line. Detrended data is then given as

$$\tilde{y}_t = y_t - \hat{\mu}^{\bar{\rho}'} z_t. \quad (2.23)$$

The test statistic is obtained by an ADF type regression without deterministic terms (2.11), based on the quasi-detrended series \tilde{y}_t and is called ADF^{GLS} . Critical values in the intercept only case are the same as for the DF test without deterministic terms and critical values for the case of a linear time trend are tabulated by Elliott et al. (1996).

An extension of the GLS quasi-differencing procedure to other unit root tests is provided by Ng and Perron (2001). They illustrate that the empirical performance of M -type unit root tests can be improved if they are constructed using GLS detrended data. Moreover, further improvements of finite sample performance can be achieved by basing the spectral density variance estimator on GLS detrended data.

2.5.2 Bootstrap unit root tests

Inference based on bootstrapped critical values can yield asymptotic refinements and offer the prospect of obtaining asymptotically correct rejection rates under the null hypothesis in cases where asymptotic approximations are non-pivotal (Horowitz, 2001). As it is shown by Basawa et al. (1991) that for $I(1)$ processes, bootstrap samples have to be constructed by imposing the unit root, all bootstrap unit root tests proposed in the literature are built upon this principle. The various approaches differ mainly with respect to the chosen resampling scheme. Several

2.5 Other approaches to unit root testing

authors suggest sieve bootstrap variants of DF and ADF tests (Chang and Park, 2003; Paparoditis and Politis, 2005; Park, 2002, 2003; Psaradakis, 2001) while others use block bootstrapping (Paparoditis and Politis, 2003) or the so-called stationary bootstrap (Swensen, 2003). A survey over the different approaches is given in Palm et al. (2008a).

Yet a different approach which appears to be well suited if the DGP is characterized by strong heteroskedasticity, such as sudden volatility breaks or stochastic volatility is the wild bootstrap (Davidson and Flachaire, 2001; Liu, 1988; Mammen, 1993). Wild bootstrap variants of M -type unit root tests have been suggested by Cavaliere and Taylor (2008, 2009). In the following, this approach is reviewed in more detail since it constitutes the basis for much of the exposition in later chapters.

2.5.2.1 Wild bootstrap unit root tests

The central feature of the wild bootstrap is that the bootstrap sample is constructed without replacement such that it replicates patterns of heteroskedasticity present in the original data. In particular, Cavaliere and Taylor (2008, 2009) propose wild bootstrap variants of the GLS detrended M -tests of Ng and Perron (2001). In the first step bootstrap residuals are generated as

$$u_t^b = \hat{e}_t w_t, \quad (2.24)$$

where \hat{e}_t are residuals from an ADF regression based on GLS detrended data and $\{w_t\}_{t=1}^T$ is an *iid* $N(0, 1)$ sequence. Hence, the only difference between the bootstrap- and regression residuals is pre-multiplied random variable w_t . Imposing the unit root, the bootstrap sample is given as

$$y_t^b = y_0^b + \sum_{i=1}^t u_i^b, \quad t = 1, \dots, T, \quad (2.25)$$

and y_0^b is set equal to y_0 . In order to ensure convergence to the same limiting distribution, the GLS detrending scheme is applied to the bootstrap sample before the M -tests (2.15)-(2.17) are subsequently computed in the usual way. However, since the bootstrap increments u_t^b are serially independent, there is no need for employing a long run variance estimator. Hence, s_{AR}^2 is replaced by a simple OLS variance estimator. The preceding steps are subsequently replicated sufficiently

2.6 Discussion of finite sample performances

often. The bootstrap critical value for the chosen nominal significance level α is then obtained as the $\alpha\%$ quantile of the resulting empirical distribution of bootstrap test statistics. It can be shown that the choice of the distribution for w_t is not restricted to the Gaussian distribution. In fact, any distribution which satisfies $E(w_t) = 0$ and $E(w_t^2) = 1$ can be chosen. Some alternative choices are given e.g. in (Davidson and Flachaire, 2001). Similarly, the construction of bootstrap residuals in (2.24) might be conducted using restricted residuals imposing the null hypothesis (Cavaliere and Taylor, 2008).

2.6 Discussion of finite sample performances

The finite sample properties of the discussed unit root tests have been studied among others by DeJong et al. (1992), Elliott et al. (1996) and Ng and Perron (2001). For a sample size $T = 100$, DeJong et al. (1992) show that rejection frequencies under the null hypothesis obtained via the semiparametric Z -tests of Phillips and Perron (1988) are characterized by substantial deviations from the nominal significance level. If the error process is driven by negative autocorrelation (AR(1) or MA(1)), the tests are substantially oversized with empirical rejection frequencies of over 60% in the case of an MA root of -0.5 . On the other hand, the Z -tests (and in particular the coefficient based test $Z_{\hat{\rho}}$) are severely undersized if the errors exhibit positive autocorrelation of either type. Due to the fact that the correct lag length $k = 1$ is chosen, the ADF test controls size rather well in the AR scenarios. Overfitting of the lag order in the AR scenarios does not adversely affect empirical size but results in a loss of power. In the cases of MA errors, the ADF test as employed by DeJong et al. (1992) with only one lagged difference yields oversized rejection frequencies for all simulated values of the MA root.

Elliott et al. (1996) demonstrate that for the MA cases, size properties of the ADF test can be substantially improved by selecting a significantly larger lag order $k = 8$ or using the Bayesian Information Criterion (BIC) to determine the appropriate lag order. Furthermore and as expected from their theoretical results, the ADF^{GLS} test offers noticeable advantages in terms of empirical power over the standard ADF test.

Ng and Perron (2001) illustrate that the choice of the lag length, and hence, the choice of the information criterion used for that purpose, can have a major impact on the empirical rejection frequencies of the considered tests. This finding is of particular importance in the case of (large) negative MA roots, where upward distorted rejection frequencies could be observed. They propose a modified Akaike information criterion (MAIC) which is more liberal than the standard AIC for error processes with large MA roots, but avoids overfitting in cases of AR or small MA roots. In a comparative simulation study, it is documented that the GLS detrended M -type tests with lag lengths chosen according to the MAIC avoid the oversizing observed for their ADF counterparts. Contrary, the M -type tests tend to underreject the null hypothesis in scenarios of strong negative serial correlation (MA or AR) with least size distortions for the statistic which employs GLS detrending in the construction of the spectral density variance estimator. In terms of size-adjusted local power, it turns out that the ADF^{GLS} test tends to be preferable, however, its severe size distortion in particular scenarios might limit its applicability in some cases of applied research.

As shown by Breitung (2002), the oversizing observed for many tests in the case of a large negative MA root also carries over to the nonparametric $J_2(m)$ statistic of Park (1990). However, it achieves relatively efficient size control if y_t follows some nonlinear process. Cavaliere and Taylor (2009) demonstrate that a recolored variant of the bootstrap M -tests is able to alleviate the undersizing of the M -tests observed for large negative MA roots. Furthermore, Cavaliere and Taylor (2008) show that the bootstrap M -tests are robust under a variety of scenarios modeling nonstationary volatility.

2.7 Conclusions

This chapter provides a basic overview over the field of unit root testing. Starting from the parametric DF tests based on the first order autoregression, semiparametric alternatives as well as fully nonparametric and bootstrap unit root tests are reviewed. Noting that correct approximation of residual serial dependence is a core issue in unit root testing, the class of M -tests based on GLS-detrended data (Ng and Perron, 2001) achieve most precise rejection frequencies under the null

hypothesis compared with ADF or Z -test variants. A recolored bootstrap variant of the M -tests (Cavaliere and Taylor, 2009) is able to further improve finite sample performance, however, comes at the cost of an increased computational burden. From the above discussion it is obvious that there is still scope for the development of further unit root tests. Conditional on good size control in cases of serially correlated error terms, increased robustness against various deviations from standard assumptions as well as reduced sensitivity with respect to the lag length selection are desirable properties for such alternative test procedures.

Chapter 3

A New Approach To Unit Root Testing

3.1 Introduction

Since the work of Granger and Newbold (1974) it is known that spurious correlations may arise if a least squares regression is fitted to uncorrelated time series which are integrated of order one. To avoid this, separating between stationary and integrated series by means of unit root tests has become a central aspect of time series econometrics. Dickey and Fuller (1979) (DF henceforth) show that for $I(1)$ processes, the t -ratio from a first order autoregression converges to a nonstandard limiting distribution which can be expressed as a functional of a Brownian motion. Accordingly, the DF unit root test is conducted by comparing this t -ratio with simulated critical values from the limiting distribution. Since then, the literature on unit root testing has been rapidly expanding. Major issues involve coping with residual autocorrelation (Said and Dickey, 1984, Phillips and Perron, 1988) and improving the power features of the tests (e.g. Elliott et al., 1996). An alternative approach to unit root testing has been proposed by Stock (1999)¹. Instead of directly testing the value of the autoregressive parameter, the so-called class of M -type tests exploits the fact that the sum of squares of an integrated process is of higher order in probability than the sum of squares of a stationary process. Perron and Ng (1996) and Ng and Perron (2001) suggest modified variants of the M tests which perform well under the null hypothesis

¹The paper dates back to 1990.

in terms of small errors in rejection probability under general forms of residual autocorrelation whilst retaining good power properties. Fully Nonparametric approaches to unit root testing which are robust against violations of standard assumptions have been proposed e.g. by Breitung and Gouriéroux (1997) and Aparico et al. (2006).

Another nonparametric approach to unit root and cointegration testing is pursued by Park and Choi (1988), Park et al. (1988) and Park (1990). The tests are based upon an appropriately scaled Wald statistic obtained from a regression of the data on a matrix of deterministic terms and superfluously included computer generated random walks. Under the unit root hypothesis this statistic converges to a nondegenerated, nuisance parameter free limiting distribution and tends to zero under the alternative hypothesis. In this chapter, a simulation based unit root test is proposed which is similar to the latter approach in that it utilizes spurious regressions for unit root testing. However, unlike in Park (1990) where the test statistic is based on the residual sum of squares of the restricted and unrestricted regressions, the proposed test statistic makes use of the different dispersion of the regression coefficient under the null and alternative hypothesis, respectively. Noting that the parameter from a spurious regression converges to a nondegenerated limiting distribution (Phillips, 1986), the parameter of an unbalanced regression of an $I(0)$ variable on an $I(1)$ regressor can be shown to be of order in probability $O_p(T^{-1})$. It is demonstrated that a consistent unit root test can be based on this distinction. In particular, regressing the appropriately scaled data sufficiently often on a random walk controlled by the analyst yields a sample of random variables which are of different order in probability under the null and alternative hypothesis, respectively. Viable test statistics can then be constructed from ranges of that random variable, which have a nondegenerated distribution under H_0 but degenerate to a one point distribution under H_1 . A simulation study is conducted to assess the empirical properties of the proposed procedure. To preview the results, it turns out that the simulation based testing approach on average offers most precise size estimates compared with ADF- and M -type tests. In large samples, the proposed test achieves higher local power than the standard ADF test but is outperformed by the ADF-GLS test proposed by Elliott et al. (1996) and the modified M -type test of Ng and Perron (2001). However, there are finite sample scenarios with residual autocorrelation where

3.2 The simulation based unit root test

the proposed test is the most powerful among those tests that are characterized by correct empirical rejection frequencies under H_0 . As an empirical illustration long run PPP among a sample of G7 economies is reconsidered. The empirical example mirrors some central results obtained in the Monte Carlo exercise. In line with the existing literature, evidence on prevalence of long run PPP is at best mixed.

3.2 The simulation based unit root test

3.2.1 The testing principle

Consider testing for a unit root in the time series $\{y_t\}_{t=1}^T$, generated as

$$y_t = d_t + x_t, \quad x_t = \rho x_{t-1} + u_t, \quad t = 1, \dots, T, \quad (3.1)$$

where $d_t = \mu' z_t$ and z_t is a vector of deterministic components and μ is a parameter vector. Regarding the sequence of innovations $\{u_t\}_{t=1}^T$ in (2.1), the following set of (weak) assumptions taken from Phillips (1986) are supposed to hold.

Assumption 3.1 (\mathcal{A}_1)

- (i) $E(u_t) = 0$ for all t ;
- (ii) $\sup_t E|u_t|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$;
- (iii) as $T \rightarrow \infty$, $\sigma^2 = \lim E(T^{-1} S_T^2)$ exists and $\sigma^2 > 0$, where $S_T = \sum_{t=1}^T u_t$;
- (iv) $\{u_t\}$ is strong mixing with mixing coefficients α_m that satisfy $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$.

Assumption $\mathcal{A}_1(ii)$ ensures the existence of the variance of u_t , denoted σ_u^2 , and thus limits the allowable heterogeneity of the process. The existence of the long run (or average) variance of the partial sum process S_T , σ^2 , is defined in $\mathcal{A}_1(iii)$. If u_t is weakly stationary, then $\sigma^2 = E(u_1^2) + 2 \sum_{k=2}^{\infty} E(u_1 u_k)$. Note that for *iid* innovations $u_t \sim iid(0, \sigma_u^2)$, the long run variance is given as $\sigma^2 = \sigma_u^2$. For further discussion of the implications of Assumption 1, see Phillips (1986). Since the work of Granger and Newbold (1974) and Phillips (1986) it is well known that the OLS estimator $\hat{\beta} = (\sum_{t=1}^T \tilde{x}_t^2)^{-1} \sum_{t=1}^T \tilde{x}_t y_t$ from the regression

$$y_t = \beta \tilde{x}_t + \varepsilon_t, \quad (3.2)$$

3.2 The simulation based unit root test

is not consistent for the true parameter $\beta = 0$ if y_t and \tilde{x}_t are two uncorrelated random walks. To make statements about the limiting distribution of $\hat{\beta}$, we assume that y_t follows (2.1) with $\rho = 1$ and $d_t = 0$ for all t , and \tilde{x}_t is generated as

$$\tilde{x}_t = \tilde{x}_{t-1} + \nu_t \quad (3.3)$$

with white noise error terms $\nu_t \sim iid N(0, \sigma_\nu^2)$ and initial condition $x_0 = 0$. Moreover, it is assumed that the innovation sequences $\{u_t\}_{t=1}^T$ and $\{\nu_t\}_{t=1}^T$ are independent. Under the unit root null hypothesis, it follows from the results in Phillips (1986) that the limiting distribution of $\hat{\beta}$ is given by

$$\hat{\beta} \xrightarrow{d} \sigma_\nu / \sigma \left[\int_0^1 W_x(a)^2 da \right]^{-1} \int_0^1 W_x(a) W_y(a) da, \quad (3.4)$$

where $W_{\tilde{x}}$ and W_y are uncorrelated standard Brownian motions and σ^2 is the long run variance of y_t as defined in $\mathcal{A}_1(ii)$. If regression (3.2) contains an intercept or linear time trend, the limiting distribution is given as a functional of demeaned or detrended Brownian motion, respectively. Park (1990) suggests a Wald statistic based on regression (3.2) to test the unit root hypothesis. The test utilizes the spurious fit obtained under H_0 for the unrestricted model (including the superfluous random walk \tilde{x}_t) whereas under H_1 , regression (3.2) does not indicate superior fit compared to the restricted model (in this case $\sum_{t=1}^T y_t^2$). The test proposed herein is similar to the approach of Park (1990) as it builds upon a regression of the data on computer generated random walks. However, the test statistic is based on the dispersion of obtained regression coefficients rather than on the spurious fit of the model. In particular, the result in (3.4) implies that $\hat{\beta} \sim O_p(1)$. If, under the alternative hypothesis, $\{y_t\}_{t=1}^T$ is a weakly stationary process with $|\rho| < 1$, it follows that $\hat{\beta}$ converges in probability to 0 at the rate of T and, hence, $\hat{\beta} \sim O_p(T^{-1})$. The testing idea is to exploit these distinct orders in probability under H_0 and H_1 , respectively.

3.2.2 Construction of the test statistic

The non degenerated limiting distribution of $\hat{\beta}$ in the unit root case implies a finite range for a sample of R realizations of $\{\hat{\beta}_r\}_{r=1}^R$. In the stationary case on the other hand, the range of the degenerated distribution of $\{\hat{\beta}_r\}_{r=1}^R$ becomes arbitrarily small for $T \rightarrow \infty$. Hence, a test statistic could be constructed as the

3.2 The simulation based unit root test

range of the sequence of regression coefficients $\{\widehat{\beta}_r\}_{r=1}^R$, obtained from regressing the data sufficiently often, say R times, on simulated random walks that are controlled by the econometrician, i.e.

$$y_t = \beta^{(r)} \widetilde{x}_t^{(r)} + \varepsilon_t^{(r)}, \quad t = 1, \dots, T, \quad r = 1, \dots, R, \quad (3.5)$$

with $\widetilde{x}_t^{(r)}$ as defined in (3.3). The range is given as,

$$\text{range}\{\widehat{\beta}^{(r)}\}_{r=1}^R = \max\{\widehat{\beta}^{(r)}\}_{r=1}^R - \min\{\widehat{\beta}^{(r)}\}_{r=1}^R. \quad (3.6)$$

Asymptotically however, the range depends on the choice of R since it is solely based on the two extremal points of the distribution of $\widehat{\beta}^{(r)}$. The likelihood of observing larger (smaller) values for the maximum (minimum) of $\widehat{\beta}^{(r)}$ increases with the sample size, R . Moreover, the realizations of the sample maximum and minimum are extremely sensitive to outliers. Unlike the extremal values, the quantiles of a given distribution function do not depend on the sample size and are more robust with respect to outliers. Hence, an alternative approach would be to use the $(1 - \alpha)\%$ interquantile range of $\{\widehat{\beta}^{(r)}\}_{r=1}^R$ as a testing device. The test statistic is then given as

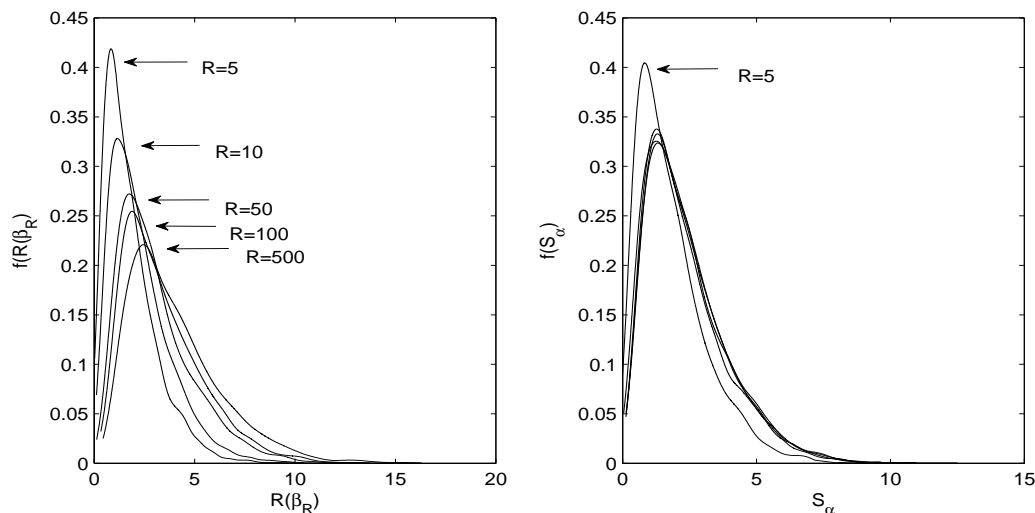
$$J_\alpha = \Xi(1 - \alpha/2) - \Xi(\alpha/2), \quad (3.7)$$

where Ξ denotes the empirical cumulative distribution function of $\{\widehat{\beta}^{(r)}\}_{r=1}^R$ and α is the chosen percentile. For instance, by setting $\alpha = 0.1$, J_α is the 90% interquantile range of $\{\widehat{\beta}^{(r)}\}_{r=1}^R$. Since the particular realizations of $\Xi(1 - \alpha/2)$ and $\Xi(\alpha/2)$ are random variables, so is J_α and its limiting distribution is given as a function of the underlying distribution function of $\widehat{\beta}^{(r)}$, Ξ , and the chosen quantile α . Figure 3.1 displays the limiting distribution of the range of $\{\widehat{\beta}^{(r)}\}_{r=1}^R$ and $J_{0.1}$ for different values of R , where the empirical distribution functions are based on 100000 replications. In both cases, innovation variances $\sigma_u^2 = \sigma_v^2 = 1$. The time dimension is chosen reasonably large with $T = 1000$ in order to ensure convergence of the estimated regression coefficients to their asymptotic distribution. However, unreported results show that a very similar picture emerges even for a time dimension as small as $T = 50$.

The left hand side graph of Figure 3.1 confirms that the limiting distribution of the range of $\{\widehat{\beta}^{(r)}\}_{r=1}^R$ depends on R . The variance of the distribution increases with R and the mode of the distribution shifts to the right as R increases. On the

3.2 The simulation based unit root test

Figure 3.1: Empirical PDFs conditional on $R, T = 1000$



Notes: Empirical pdfs are approximated by a Gaussian kernel density estimator with bandwidth chosen according to Silverman's rule of thumb.

other hand, as shown in the right hand side graph, the limiting distribution of $J_{0,1}$ displays convergence to a unique distribution even for relatively small values of R .

To ensure that the limiting distribution of J_α is free of nuisance parameters, σ and σ_ν have to be removed from the limiting distribution of $\widehat{\beta}^{(r)}$. While σ_ν is under control of the econometrician and can thus be set to unity, it is necessary to scale y_t by a consistent estimate of the long run standard deviation, σ , in order to obtain a nuisance free limiting distribution. This issue will be considered in detail in the following section. Finally, it has to be noted that, as J_α depends on the random draws $\{\nu_t\}_{t=1}^T$, the test decision also depends on these random draws and might vary for different draws of $\{\nu_t\}_{t=1}^T$. This feature of additional randomness is common for all simulation based test approaches, as, for example, bootstrap tests or the variable addition unit root test of Park (1990).

3.2.3 Consistent estimation of the long run variance

As outlined in Section 3.2.1, a consistent estimate of the long run variance σ^2 is required to implement the proposed unit root test under general forms of serial dependence. It is shown in Phillips (1987) that robust estimators of the

3.2 The simulation based unit root test

asymptotic covariance in the spirit of White (2000) or Newey and West (1987) are feasible choices. Phillips (1987) and Phillips and Perron (1988) advocate the use of the Newey and West (1987) procedure (see Chapter 2.4 for details). They show that for any weakly stationary innovation sequence, the Newey-West estimator is equivalent to 2π times the Bartlett estimate of the spectral density at frequency zero. Since the latter estimate is always nonnegative, it follows that the Newey-West estimator also ensures a nonnegative estimate of the long run variance. However, it turns out that unit root tests based on Newey and West (1987) type variance estimators often perform poorly in finite samples with serially correlated innovations. An alternative long run variance estimator which will be applied in the construction of the J_α statistic has been proposed by Perron and Ng (1996). As outlined in Chapter 2.4.2, Perron and Ng (1996) propose to use an autoregressive spectral density estimator at frequency zero of u_t , denoted s_{AR}^2 which is based upon an auxiliary ADF regression with previously determined lag order. To recall the exposition in Chapter 2.4.2, the estimator s_{AR}^2 is then constructed as

$$s_{AR}^2 = \hat{\sigma}_{ek}^2 / (1 - \hat{\beta}(1))^2, \quad (3.8)$$

with $\hat{\sigma}_{ek}^2 = (T-k)^{-1} \sum_{t=k+1}^T \hat{e}_{tk}^2$, $\hat{\beta}(1) = \sum_{j=1}^k \hat{\beta}_j$, where $\hat{\beta}_i$ and $\{\hat{e}_{tk}\}$ are obtained from the auxiliary ADF regression. Ng and Perron (2001) demonstrate that the finite sample size of the M -tests can be further improved by estimating s_{AR}^2 from an ADF regression based on GLS detrended data (see Chapter 2.5.1). This approach will be adopted in the construction of the J_α statistic in the following.

3.2.3.1 Lag length selection

It is long recognized that the appropriate choice of the truncation lag k is important for the performance of unit root tests based either on parametric approximations of the error terms' serial dependence (Said and Dickey, 1984) or tests based on robust variance estimators (Ng and Perron, 2001; Phillips and Perron, 1988). While underfitting of k often leads to severe size distortions, too liberal a choice of k might adversely affect the power of the tests. Ng and Perron (2001) point out that even the comparatively liberal Akaike Information Criterion (AIC) tends to select too small a lag order if there is a negative moving average root in the innovation process. Therefore, they propose the *modified* AIC (MAIC). It is

3.2 The simulation based unit root test

given as

$$\text{MAIC}(k) = \ln(\widehat{\sigma}_k^2) + \frac{2(\tau_T(k) + k)}{T - k_{max}}, \quad (3.9)$$

where $\tau_T(k) = (\widehat{\sigma}_{ek}^2)^{-1} \widehat{\beta}_0^2 \sum_{t=k_{max}+1}^T y_{t-1}^2$ and $\widehat{\sigma}_{ek}^2 = (T - k_{max})^{-1} \sum_{t=k_{max}+1}^T \widehat{e}_{tk}^2$. The chosen lag length is then obtained as $k_{MAIC} = \arg \min_k(\text{MAIC})$ with admissible values of $k \in [0, k_{max}]$, where k_{max} is usually given by the integer part of $12(T/100)^{1/4}$. While Ng and Perron (2001) suggest to base the MAIC on an ADF regression based on GLS detrended data, Perron and Qu (2007) recommend to base the information criterion on a standard ADF regression while still employing GLS detrending for the construction of the spectral density variance estimator and the test statistics. They argue that this hybrid procedure improves the small sample properties of the considered tests.

3.2.4 Deterministic terms

There are two viable approaches of accounting for deterministic terms in the DGP: OLS or (quasi-)GLS detrending. While the first proceeds by simply introducing a vector of deterministic terms in the test regression 3.5, GLS detrending as proposed by Elliott et al. (1996) is reviewed in Chapter 2.5.1. Elliott et al. (1996) and Ng and Perron (2001) have shown that significant power gains can be obtained by basing ADF or M -type unit root tests test on GLS detrended data. In the following, both approaches implemented and compared by means of the Monte Carlo study.

3.2.5 Critical values

Critical values for two variants of the $J_{0.1}$ statistic are documented in Table 3.1. The implementations of the statistic differ with respect to the treatment of the deterministic terms. Values underneath $J_{0.1}^{Reg}$ refer to the implementation of the statistic where deterministic terms are included in the test regression, while critical values for the statistic based on GLS detrending are headed by $J_{0.1}^{GLS}$. All critical values are generated using 100000 Monte Carlo replications and setting $\sigma_\nu = \sigma_u = 1$, $\alpha = 10\%$ and $R = 50$. The employed spectral density variance estimator is based on GLS detrended data as discussed in Section 3.2.3 where the correct lag length $k = 0$ is selected.

3.3 Finite sample properties

Table 3.1: Critical values, $J_{0.1}$

T	$J_{0.1}^{Reg}$			$J_{0.1}^{GLS}$			$J_{0.1}^{Reg}$			$J_{0.1}^{GLS}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
	Intercept						Trend					
25	0.50	0.69	0.82	0.34	0.50	0.61	0.59	0.75	0.84	0.21	0.28	0.33
50	0.42	0.61	0.75	0.34	0.51	0.63	0.46	0.61	0.71	0.19	0.26	0.32
100	0.40	0.58	0.72	0.33	0.53	0.68	0.41	0.56	0.65	0.18	0.26	0.32
250	0.38	0.57	0.70	0.35	0.56	0.72	0.39	0.53	0.63	0.19	0.27	0.33
500	0.38	0.57	0.70	0.35	0.57	0.75	0.39	0.53	0.63	0.19	0.27	0.34
1000	0.38	0.56	0.70	0.36	0.58	0.76	0.38	0.52	0.62	0.19	0.28	0.34
10000	0.38	0.57	0.70	0.36	0.60	0.78	0.38	0.52	0.62	0.19	0.28	0.34

Notes: Data are generated according to equations (3.1) and (3.3) with $d_t = 0$ and $u_t, \nu_t \sim iid N(0, 1)$. Results are based on 100000 replications and $R = 50$. The variance estimator s_{AR}^2 is constructed with $k = 0$ based on GLS detrended data.

3.3 Finite sample properties

The finite sample properties of both implementations of the J_α statistic are analyzed by means of a Monte Carlo study. Data is generated according to model (3.1) for $t = -49, \dots, -1, 0, 1, \dots, T$ and the pre-sample values are discarded. Besides the benchmark scenario given by $u_t \sim iid N(0, 1)$, serially dependent innovation processes formalized by means of first order moving average

$$\text{MA}(1): \quad u_t = \Theta e_{t-1} + e_t, \quad e_t \sim iid N(0, 1), \quad (3.10)$$

and first order autoregressive innovation structures

$$\text{AR}(1): \quad u_t = \varrho u_{t-1} + e_t, \quad e_t \sim iid N(0, 1), \quad (3.11)$$

are considered. To capture a wide range of correlation patterns, both cases are simulated for parameter values $\Theta, \varrho \in \{-0.8, -0.5, 0.5, 0.8\}$. The random walk \tilde{x}_t , needed for the construction of J_α is generated as a Gaussian random walk according to (3.3) with starting value $\tilde{x}_0 = 0$. The relative performance of the proposed simulation based unit root test is assessed by comparing it with the standard ADF test, the ADF-GLS test by Elliott et al. (1996), the \bar{M}^{GLS} test proposed by Ng and Perron (2001) and, finally, with the nonparametric J_2 test proposed by Park (1990). All of these benchmark tests are discussed in the previous chapter. In line with Breitung (2002), a four dimensional unit root process

is employed in the construction of the J_2 statistic. The lag length is selected for all tests according to the suggestion of Perron and Qu (2007) by the MAIC based on a standard ADF regression. Empirical size is evaluated under the null hypothesis of $\rho = 1$ and the nominal significance level is 5%, however, results remain qualitatively unchanged if alternative nominal significance levels are chosen. The number of replications is set to 5000. Empirical size estimates are based on simulated critical values for all tests since exact critical values are not tabulated in the literature for some of the tests. Exact critical values are generated from 100000 replications of model (3.1) under the null hypothesis with white noise error terms. Rejection frequencies under H_1 are calculated for the local alternative $H_1 : \rho = 1 + \bar{c}/T$, where $\bar{c} = -7$ and $\bar{c} = -13.5$ in the intercept and trend case, respectively. Throughout, rejection frequencies under H_1 are size adjusted, such that the reported local power results are based on test specific nominal levels which ensure a 5% rejection frequency under H_0 . Tables 3.2 and 3.3 list rejection frequencies under the null hypothesis and bold entries highlight rejection frequencies which are not covered by the 95% confidence interval [4.4, 5.6] constructed around the nominal level $\alpha = 0.05$ as $\alpha \pm 1.96\sqrt{\alpha(1-\alpha)/5000}$. Tables 3.4 and 3.5 document size adjusted local power results. Italic entries denote power estimates which rely on substantial size adjustments as the corresponding rejection frequencies under H_0 are outside the 95% confidence interval [4.4, 5.6].

3.3.1 Rejection frequencies in the unit root case

Rejection frequencies obtained for white noise innovations illustrate that the proposed test achieves a high degree of size control. Irrespective of the maintained deterministic components both variants of the J_α test achieve more robust size control than the ADF or M -type benchmark tests. Only one significant deviation from the nominal significance level can be observed for each variant of the J_α test in the intercept only case whereas the ADF and ADF-GLS statistics yield four significant deviations from the nominal significance level. For the \bar{M}^{GLS} and the $J_2(4)$ statistics three, respectively two significant deviations can be observed. In the case of a maintained time trend, the general picture is very similar with two significant violations of the 5% significance level obtained for $J_{0,1}^{Reg}$ while rejection frequencies for $J_{0,1}^{GLS}$ vary insignificantly around the nominal level. Except for the $J_2(4)$ statistic, all implemented statistics tend to display empirical rejection

3.3 Finite sample properties

frequencies which are below the nominal significance level. Largest (downward) violations of the 5% level are observed for the ADF-GLS statistic with empirical rejection frequencies as low as 3.2% for $T = 50$. The observed downward distortions of rejection frequencies are presumably induced by spuriously included lags in the test regressions or in the construction of the spectral density variance estimator. Size distortions are generally more pronounced for small sample sizes which is in line with Cheung and Lai (1995) who demonstrate that the critical values of the ADF statistic exhibit a nonlinear dependence on the chosen lag length k which vanishes for increasing T . If the simulations are based on the correct lag length (i.e. $k = 0$, unreported), the observed downward bias of empirical rejection frequencies disappears. The nonparametric $J_2(4)$ statistic does not show systematic underrejections and displays two significant violations of the nominal level in both scenarios.

If the random walk innovations are generated by an MA structure with negative coefficients, rejection frequencies are much less precise for all statistics than under uncorrelated innovations. Especially for (large) negative MA coefficients ($\Theta = -0.8$) and $T < 100$, all considered statistics display errors in rejection probabilities of magnitudes which appear prohibitive for the application of the statistics in empirical analyses. Considering larger sample sizes, the documented oversizing is somewhat reduced with similar empirical rejection frequencies for the $J_{0.1}^{GLS}$ and both ADF variants, ranging around the six percent mark for $T = 500$. The \bar{M}^G statistic achieves relatively good size control for $T = 100$ but tends to underreject H_0 in larger samples, a finding which is in line with results reported in Ng and Perron (2001). Least precise empirical rejection frequencies are obtained by the $J_2(4)$ statistic with double-digit rejection frequencies even for samples as large as $T = 500$. In the case of an included time trend, observed size distortions are even more pronounced for most statistics with the ADF variants achieving at least some degree of size control (rejection frequencies between 6.2% and 7.1%) for sample sizes of $T \geq 250$.

For moderate negative MA dynamics ($\Theta = -0.5$) size distortions are much less pronounced. Most reliable size estimates are obtained via both variants of the ADF test and by the \bar{M}^{GLS} statistic. For $T = 100$, empirical rejection frequencies vary between 5.1% and 5.8% in the intercept only case and between 5.9% and 6.8% in the case including a time trend. Rejection frequencies obtained by $J_{0.1}^{Reg}$

3.3 Finite sample properties

Table 3.2: Empirical size

T	Intercept						Trend					
	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$
$u_t \sim iid N(0, 1)$												
25	5.3	5.1	4.2	3.8	5.6	5.4	5.4	5.4	4.2	3.9	5.7	5.6
50	4.6	4.7	4.1	3.7	4.0	5.1	4.5	5.4	3.2	3.3	3.8	5.1
100	4.3	4.9	4.2	4.1	4.2	5.0	4.1	4.7	3.3	3.3	3.6	5.1
250	4.8	4.8	4.4	4.4	4.7	5.8	3.8	4.7	3.4	3.5	3.7	4.8
500	4.4	4.1	4.3	3.9	4.0	4.9	4.8	4.7	4.2	4.3	4.5	6.5
1000	4.9	5.0	4.8	4.6	4.7	3.6	5.1	5.0	5.1	4.6	4.7	3.6
MA case: $\Theta = -0.8$												
25	44.1	41.5	34.4	38.1	38.9	46.0	56.1	46.7	48.9	51.2	52.1	31.6
50	28.5	26.7	20.0	21.9	19.0	49.3	49.5	45.3	35.3	36.8	36.6	48.4
100	16.0	13.6	9.0	10.1	5.7	38.8	29.7	23.5	18.1	17.6	14.1	49.8
250	9.5	8.2	6.7	7.4	2.3	23.0	13.7	11.2	6.9	6.8	2.1	30.5
500	6.8	6.3	6.0	5.8	2.5	14.4	10.5	8.4	6.7	6.2	1.2	23.3
1000	7.0	6.3	6.4	6.7	3.6	6.8	8.6	7.1	7.4	6.7	1.8	8.6
MA case: $\Theta = -0.5$												
25	17.0	16.2	13.5	16.2	17.7	15.2	25.7	20.7	22.6	24.6	28.2	14.4
50	10.4	10.1	6.9	7.3	7.5	12.6	18.1	15.8	11.0	11.8	13.7	13.8
100	7.3	7.4	5.1	5.8	5.1	8.7	10.9	9.2	5.9	6.6	6.8	10.8
250	6.1	6.2	5.5	5.3	4.9	7.4	6.1	6.6	4.5	4.6	4.2	7.2
500	5.1	4.9	4.9	4.7	4.7	5.9	6.5	6.1	5.4	4.7	4.4	7.7
1000	5.4	5.3	5.5	5.3	5.4	4.0	6.1	5.8	5.7	5.6	5.5	4.2
MA case: $\Theta = 0.5$												
25	3.7	3.6	3.1	0.8	6.3	4.1	2.6	2.4	1.4	0.4	5.8	3.7
50	3.6	4.8	2.6	1.4	4.7	4.3	2.0	3.4	1.1	0.4	4.8	3.8
100	4.4	5.1	3.4	3.2	5.0	4.5	4.0	4.9	2.3	2.3	5.7	4.4
250	4.3	4.9	4.0	4.0	4.5	5.6	3.7	4.8	2.5	2.6	4.3	4.6
500	4.2	4.2	4.0	3.7	4.0	4.8	4.7	5.2	3.7	3.5	4.6	6.3
1000	4.8	4.9	4.8	4.3	4.8	3.6	5.3	5.1	4.6	4.6	5.3	3.5
MA case: $\Theta = 0.8$												
25	4.9	4.7	2.8	0.7	9.6	4.0	3.5	3.1	1.5	0.1	9.1	3.5
50	4.4	5.3	2.8	1.4	6.7	4.1	3.6	4.4	1.2	0.4	8.6	3.7
100	4.3	5.6	2.8	2.3	6.0	4.4	3.9	4.8	1.5	0.9	6.5	4.4
250	4.6	5.3	3.8	3.8	5.7	5.6	4.1	5.6	2.1	1.7	5.8	4.6
500	4.4	4.6	4.0	3.4	4.6	4.8	5.6	5.9	3.6	3.4	6.2	6.2
1000	4.9	4.8	4.6	4.3	5.1	3.6	5.6	5.7	4.7	4.4	6.2	3.5

Notes: $J_{0.1}^{Reg}$ denotes the simulation based test including deterministic terms in the test regressions while $J_{0.1}^{GLS}$ refers to the test based on GLS detrended data. ADF denotes the augmented Dickey-Fuller statistic, ADF^G denotes the ADF-GLS test of Elliott et al. (1996) and \bar{M}^G the M -type test of Ng and Perron (2001) based on GLS detrending. Finally, $J_2(4)$ refers to the variable addition test proposed by Park (1990) where the dimension of the multivariate random walk is equal to 4. To facilitate interpretation of the Tables, bold entries indicate rejection frequencies which are not covered by the 95% confidence interval [4.4%, 5.6%] around the nominal 5% level constructed as $\alpha \pm 1.96\sqrt{\alpha(1-\alpha)}/5000$, $\alpha = 0.05$. Rejection frequencies under the null hypothesis are calculated for data generated according to model (3.1) with $d_t = 0$ and $\rho = 1$. MA and AR error processes are generated by (3.10) and (3.11), respectively. 5000 replications are generated throughout and test statistics $J_{0.1}^{Reg}$ and $J_{0.1}^{GLS}$ are based on $R = 50$. For all statistics, the lag length is chosen according to the $MAIC$ applied to OLS demeaned or detrended data.

3.3 Finite sample properties

and $J_{0.1}^{GLS}$ are slightly more liberal in small samples. For example, for $T = 100$, 7.3% and 7.4% rejections of H_0 are obtained in the intercept case and 10.9% and 9.2% in the trend case, respectively.

Positive MA dynamics appear to induce empirical rejection frequencies of less than the nominal level for most of the employed statistics. The notable exception is the \bar{M}^{GLS} statistic which, in contrast, tends to overreject H_0 . The latter feature is most pronounced for $\Theta = 0.8$ with significant deviations from the nominal 5% level even for large time dimensions. The simulation based statistics $J_{0.1}^{Reg}$ and $J_{0.1}^{GLS}$ display a rather robust performance. Downward violations are less frequent and much less pronounced than observed for both implementations of the ADF statistic. For these, substantially downward biased rejection frequencies (often below 2%, in some cases even of less than 1%) imply large type two errors in applied work. Hence, in scenarios with positive serial correlation of the MA type and a small to moderate sample size, these tests might often lack the ability to reject the unit root hypothesis, even if the series under investigation is indeed stationary.

If the random walk innovations are negatively serially correlated by means of an AR(1) process (upper half of Table 3.3), the J_α statistics outperform the benchmark tests in most instances. In the intercept only case with negative AR coefficients, results for the ADF and M-type statistics indicate a tendency of underrejecting H_0 , mostly by a significant margin. For instance consider a time dimension of $T = 100$. Rejection frequencies of 0.8% ($\varrho = -0.8$) and 3.3% ($\varrho = -0.5$) are obtained for the \bar{M}^{GLS} statistic. Slightly less pronounced underrejections are reported for both variants of the ADF statistic. Interestingly, some significant oversizing can be observed for the nonparametric alternative $J_2(4)$. In contrast, both $J_{0.1}^{Reg}$ and $J_{0.1}^{GLS}$ yield rejection frequencies close to the nominal level for time dimensions $T \geq 50$ in the intercept only case and for $T \geq 100$ if a time trend is included. Moreover, only marginal differences are visible among the two alternatives with slightly more accurate rejection frequencies documented for $J_{0.1}^{GLS}$.

If the AR coefficient ϱ is positive, rejection frequencies for both variants of the ADF statistic remain too low for most combinations of T and ϱ . In many cases, rejection frequencies are around or below 3.5% for reasonably large time dimensions such as $T \in [100, 250]$. A notable exception is given by the standard ADF

3.3 Finite sample properties

Table 3.3: Empirical size, continued

T	Intercept						Trend					
	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$
AR case: $\rho = -0.8$												
25	7.2	6.9	4.1	3.6	2.9	21.5	10.8	9.8	6.3	5.7	3.5	24.8
50	5.1	5.0	3.0	2.5	0.4	15.3	7.0	6.1	3.0	2.3	0.3	18.8
100	4.4	4.8	3.6	3.2	0.8	9.8	5.0	4.4	2.6	2.4	0.2	12.2
250	4.7	4.7	4.3	3.9	2.5	7.6	4.1	4.8	3.0	3.2	0.9	7.6
500	4.5	4.2	4.2	3.8	3.4	6.0	5.3	5.1	4.4	3.8	2.1	8.0
1000	4.9	4.6	4.8	4.4	4.2	4.0	5.2	4.8	5.1	4.4	3.4	4.2
AR case: $\rho = -0.5$												
25	9.1	8.2	6.9	6.6	8.2	9.9	13.3	11.8	11.3	11.0	12.6	11.0
50	5.9	5.9	3.6	3.1	3.1	7.9	7.5	7.0	3.8	3.5	3.7	8.6
100	4.8	5.4	3.7	3.4	3.3	6.3	5.1	5.2	3.1	2.8	2.5	7.1
250	4.7	4.9	4.2	4.9	4.2	6.3	4.3	4.7	3.1	3.3	3.0	5.6
500	4.5	4.4	4.3	3.9	4.0	5.3	5.2	4.9	4.6	4.0	3.9	6.8
1000	4.9	4.7	4.9	4.6	4.6	3.8	5.3	5.1	5.2	4.6	4.5	3.9
AR case: $\rho = 0.5$												
25	4.0	4.1	3.7	0.6	9.6	2.6	3.0	2.6	1.5	0.1	10.8	2.8
50	3.8	4.7	3.2	1.9	6.8	3.4	2.1	2.9	1.3	0.7	8.6	2.8
100	4.4	4.8	4.2	3.7	5.5	4.0	3.6	4.4	3.2	2.9	6.9	3.7
250	4.4	4.7	4.4	4.5	5.1	5.2	3.7	4.5	3.2	3.4	4.7	4.2
500	4.3	4.4	4.3	3.8	4.0	4.7	4.7	4.9	4.3	4.3	4.9	6.0
1000	4.9	4.7	4.9	4.7	5.0	3.6	5.0	5.0	5.0	4.7	4.9	3.5
AR case: $\rho = 0.8$												
25	13.2	12.6	5.5	2.2	29.0	1.7	14.4	10.2	3.4	0.7	38.8	1.4
50	6.3	6.2	5.1	3.6	14.2	2.1	7.4	6.7	4.0	2.8	27.3	1.7
100	4.5	5.3	4.6	4.1	8.3	2.7	3.5	4.0	4.0	3.5	11.5	2.6
250	3.9	4.7	4.8	4.4	5.2	4.2	3.1	3.7	3.5	3.4	5.6	3.3
500	4.0	4.2	4.5	4.0	4.5	4.3	4.0	4.4	4.8	4.4	5.4	5.2
1000	4.6	4.4	5.0	4.3	4.5	3.4	4.8	4.7	4.9	4.7	5.2	3.1

Notes: See Table 3.2

test in the intercept only case with $\rho = 0.8$, where no significant deviation from the nominal level can be detected. On the other hand, significant overrejections of H_0 are obtained for \bar{M}^{GLS} in samples with a small time dimension. If $\rho = 0.5$ empirical sizes of up to 8.6% (trend case) are observed for $T = 50$. If $\rho = 0.8$, size distortions are even more pronounced. In the intercept case, empirical rejection frequencies between 8.3% ($T = 100$) and 29.0% ($T = 25$) are documented, while in the trend case, rejection frequencies range between 11.5% ($T = 100$) and 38.8% ($T = 25$). As in the case of negative AR correlation, both variants of the $J_{0.1}$ statistic remain most robust in terms of empirical rejection frequencies close to

the nominal level. Significant size distortions are mostly restricted to small time dimensions and less pronounced than those obtained by the benchmark statistics in the case of moderate correlation ($\rho = 0.5$). If $\rho = 0.8$ and $T \leq 50$, upward distorted size estimates are documented for the $J_{0,1}$ statistics, however, these distortions are much smaller and vanishes faster than those observed for the \bar{M}^{GLS} statistic. For comparison, results obtained for the nonparametric statistic $J_2(4)$ indicate too low rejection frequencies for most scenarios of positive AR error processes.

To conclude this section, the conducted Monte Carlo study confirms that both variants of the simulation based J_α unit root test are very competitive in terms of their implied rejection frequencies under the unit root null hypothesis. The only exception is a scenario of (strong) negative serial correlation formalized as an MA(1) process, however, in this particular case none of the considered benchmark statistics yields fully satisfactory results either. Across all considered scenarios it turns out that the GLS detrending variant of J_α offers at most marginal improvements over the standard approach where deterministic terms are included in the test regression.

3.3.2 Size adjusted local power

Size adjusted local power estimates for *iid* and MA innovations are documented in Table 3.4. Entries in *italic* indicate that these power estimates are based on substantial size adjustment, as the corresponding rejection frequencies under H_0 are not covered by the 95% confidence interval around the nominal level. Hence, these results should be interpreted cautiously, since rejection frequencies are unreliable under H_0 .

In the intercept only case, the results document that the proposed $J_{0,1}$ statistics yield (size adjusted) local power estimates of slightly below 30% for the largest time dimension $T = 1000$ irrespective of the prevalence of serial correlation. For these large T scenarios the documented results indicate that $J_{0,1}^{GLS}$ offers some advantages in terms of size adjusted local power of around 2-3 percentage points compared with $J_{0,1}^{Reg}$. Compared with the standard ADF- or the $J_2(4)$ statistic, both $J_{0,1}$ statistics display sizeable power advantages of around 10 percentage points. However, the ADF-GLS and the \bar{M}^{GLS} tests achieve the highest (size adjusted) local power in scenarios including only an intercept term. On average

3.3 Finite sample properties

Table 3.4: Size adjusted local power

T	Intercept						Trend					
	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	<i>ADF</i>	<i>ADF^G</i>	\bar{M}^G	$J_2(4)$	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	<i>ADF</i>	<i>ADF^G</i>	\bar{M}^G	$J_2(4)$
$u_t \sim iid N(0, 1)$												
25	21.2	21.2	<i>13.5</i>	<i>24.2</i>	19.7	15.5	31.5	27.4	<i>28.0</i>	<i>35.7</i>	<i>31.5</i>	14.1
50	23.1	23.3	<i>14.5</i>	<i>26.9</i>	25.8	16.1	32.6	27.7	<i>26.7</i>	<i>35.7</i>	<i>35.9</i>	15.3
100	<i>24.4</i>	20.9	<i>15.6</i>	<i>25.8</i>	<i>25.1</i>	15.6	<i>32.2</i>	28.4	<i>26.7</i>	<i>34.9</i>	<i>35.4</i>	15.1
250	25.4	21.6	16.6	26.3	25.9	<i>14.6</i>	<i>33.3</i>	28.4	<i>29.3</i>	<i>37.8</i>	<i>38.0</i>	15.6
500	27.2	<i>25.9</i>	<i>18.0</i>	<i>33.1</i>	<i>32.8</i>	16.2	31.4	27.9	<i>27.7</i>	<i>37.6</i>	<i>38.0</i>	<i>14.8</i>
1000	27.6	28.4	17.1	34.8	34.4	<i>17.0</i>	31.3	29.5	27.3	39.7	40.1	<i>13.8</i>
MA case: $\Theta = -0.8$												
25	<i>17.5</i>	<i>13.8</i>	<i>16.8</i>	<i>17.4</i>	<i>8.3</i>	<i>13.4</i>	<i>19.5</i>	<i>15.1</i>	<i>28.7</i>	<i>28.2</i>	<i>3.8</i>	<i>12.5</i>
50	<i>21.3</i>	<i>16.5</i>	<i>19.3</i>	<i>20.0</i>	<i>19.0</i>	<i>16.7</i>	<i>26.3</i>	<i>20.2</i>	<i>29.7</i>	<i>28.7</i>	<i>13.3</i>	<i>15.2</i>
100	<i>21.5</i>	<i>17.9</i>	<i>18.1</i>	<i>19.1</i>	<i>18.6</i>	<i>20.4</i>	<i>31.2</i>	<i>24.2</i>	<i>28.7</i>	<i>28.2</i>	<i>28.0</i>	<i>19.0</i>
250	<i>22.2</i>	<i>18.0</i>	<i>15.5</i>	<i>17.8</i>	<i>15.6</i>	<i>19.5</i>	<i>29.1</i>	<i>21.2</i>	<i>23.5</i>	<i>23.0</i>	<i>20.4</i>	<i>22.3</i>
500	<i>25.9</i>	<i>21.6</i>	<i>17.0</i>	<i>25.7</i>	<i>23.1</i>	<i>20.5</i>	<i>28.3</i>	<i>22.8</i>	<i>21.7</i>	<i>25.9</i>	<i>21.5</i>	<i>22.0</i>
1000	<i>27.6</i>	<i>25.0</i>	<i>17.9</i>	<i>30.4</i>	<i>29.6</i>	<i>19.7</i>	<i>30.0</i>	<i>26.8</i>	<i>23.8</i>	<i>30.1</i>	<i>26.1</i>	<i>19.6</i>
MA case: $\Theta = -0.5$												
25	<i>22.3</i>	<i>21.1</i>	<i>19.4</i>	<i>21.6</i>	<i>19.0</i>	<i>19.4</i>	<i>25.2</i>	<i>23.4</i>	<i>32.3</i>	<i>33.6</i>	<i>9.6</i>	<i>16.1</i>
50	<i>22.0</i>	<i>19.9</i>	<i>16.5</i>	<i>20.0</i>	<i>19.3</i>	<i>18.5</i>	<i>28.8</i>	<i>26.9</i>	<i>24.7</i>	<i>25.6</i>	<i>27.1</i>	<i>19.2</i>
100	<i>21.4</i>	<i>18.5</i>	14.9	<i>20.8</i>	19.8	<i>18.2</i>	<i>27.6</i>	<i>24.2</i>	<i>21.5</i>	<i>24.8</i>	<i>23.7</i>	<i>18.7</i>
250	<i>23.5</i>	<i>19.9</i>	15.0	23.9	24.2	<i>15.5</i>	<i>32.9</i>	<i>26.7</i>	24.5	31.0	<i>29.8</i>	<i>17.6</i>
500	26.8	25.2	17.2	30.2	28.8	<i>16.9</i>	<i>30.1</i>	<i>26.7</i>	24.3	34.8	34.1	<i>15.9</i>
1000	27.3	28.4	17.3	32.9	32.9	<i>17.4</i>	<i>31.6</i>	<i>29.4</i>	<i>26.9</i>	35.6	35.9	<i>14.4</i>
MA case: $\Theta = 0.5$												
25	<i>14.3</i>	<i>14.9</i>	<i>6.1</i>	<i>19.4</i>	<i>3.7</i>	<i>13.3</i>	<i>22.3</i>	<i>17.5</i>	<i>18.8</i>	<i>16.1</i>	<i>31.5</i>	<i>2.8</i>
50	<i>17.5</i>	17.4	<i>7.4</i>	<i>20.0</i>	13.1	<i>15.3</i>	<i>17.8</i>	<i>13.4</i>	<i>16.1</i>	<i>12.8</i>	<i>23.4</i>	<i>6.2</i>
100	20.4	19.3	<i>11.3</i>	<i>22.8</i>	20.6	14.7	<i>22.2</i>	17.3	<i>19.1</i>	<i>15.5</i>	<i>21.4</i>	19.7
250	<i>24.6</i>	20.7	<i>13.2</i>	<i>24.6</i>	24.4	14.4	<i>25.4</i>	20.0	<i>25.1</i>	<i>23.9</i>	<i>32.6</i>	30.6
500	<i>27.2</i>	<i>25.5</i>	<i>16.2</i>	<i>31.1</i>	<i>30.8</i>	16.0	25.8	20.9	<i>26.4</i>	<i>24.0</i>	<i>33.5</i>	<i>33.6</i>
1000	27.4	29.2	16.6	<i>35.2</i>	34.7	<i>16.9</i>	28.5	22.5	30.1	26.7	36.5	<i>37.1</i>
MA case: $\Theta = 0.8$												
25	9.0	10.1	<i>5.6</i>	<i>15.3</i>	<i>2.1</i>	<i>13.5</i>	<i>16.5</i>	<i>13.1</i>	<i>14.6</i>	<i>11.2</i>	<i>27.1</i>	<i>1.0</i>
50	15.1	15.0	<i>6.8</i>	<i>18.4</i>	<i>9.8</i>	<i>14.9</i>	<i>15.4</i>	11.1	<i>12.9</i>	<i>7.4</i>	<i>16.9</i>	<i>3.3</i>
100	<i>19.0</i>	15.6	<i>8.8</i>	<i>19.7</i>	<i>15.6</i>	14.6	<i>21.2</i>	15.4	<i>17.1</i>	<i>14.4</i>	<i>23.7</i>	14.9
250	22.3	19.4	<i>12.0</i>	<i>22.5</i>	<i>21.3</i>	14.5	<i>22.8</i>	19.4	<i>22.7</i>	<i>18.7</i>	<i>27.1</i>	25.0
500	26.8	24.0	<i>14.1</i>	<i>31.1</i>	29.4	16.0	24.1	<i>20.9</i>	<i>24.3</i>	<i>20.8</i>	<i>28.9</i>	<i>28.7</i>
1000	27.1	29.4	15.9	<i>33.5</i>	32.0	<i>16.8</i>	26.9	<i>22.2</i>	28.2	22.8	<i>33.4</i>	<i>31.7</i>

Notes: To facilitate interpretation of the tables, italic entries indicate size adjusted power estimates, corresponding to rejection frequencies under H_0 which are not covered by the 95% confidence interval [4.4%, 5.6%]. For further notes see Table 3.2.

both tests offer a positive power differential compared with the $J_{0.1}$ statistics of around 6 percentage points in the large samples ($T = 1000$). In smaller samples however, this advantage is less pronounced.

3.3 Finite sample properties

Serially correlated innovations reduce local power estimates of all tests for small time dimensions. However, it is noteworthy that the $J_{0.1}$ tests appear to be less affected by this adverse effect compared with the benchmark statistics. Consider, for instance, the MA case with $\Theta = 0.8$ and $T = 50$. In this scenario, the $J_{0.1}$ statistics yield about 50% higher rejection frequencies than the \bar{M}^{GLS} statistic which appears to be most affected by the serial correlation induced power loss in small samples.

If tests are implemented to account for a linear time trend under the alternative hypothesis, the ADF-GLS and \bar{M}^{GLS} statistics remain most powerful in large samples. However, the power differential compared with the standard ADF test is less pronounced as in the intercept case, resembling a result of Elliott et al. (1996). In contrast to the intercept case, the $J_{0.1}$ statistics are no longer superior to the standard ADF statistic, however, generally outperform the $J_2(4)$ statistic. It is noteworthy, that GLS detrending does obviously not improve power features of the $J_{0.1}$ test in the trend case. On the contrary, $J_{0.1}^{Reg}$ is more powerful than $J_{0.1}^{GLS}$ in most instances. As before, residual serial correlation reduces local power estimates in small samples.

Table 3.5 lists local power estimates for data generated with AR(1) innovations. The most notable differences compared with the case of MA innovations can be observed for small sample ($T \in [25, 50]$) scenarios with positive AR coefficients where size adjusted rejection frequencies are substantially depressed, often to a degree where the tests become biased. For large time dimensions, the main conclusions drawn from the results in Table 3.4 persist.

To summarize local power estimates, it turns out that for large sample sizes the ADF-GLS and \bar{M}^{GLS} are the most powerful among the considered tests. Moreover, unlike for the ADF statistics, GLS detrending offers only minor improvement of local power estimates among the $J_{0.1}$ statistics in the case of an included intercept term, while in the scenarios additionally featuring a linear time trend, $J_{0.1}^{GLS}$ does no longer yield any improvements of size adjusted local power results.

3.4 Empirical illustration: PPP of G7 economies

Table 3.5: Size adjusted local power, continued

T	Intercept						Trend					
	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$
AR case: $\rho = -0.8$												
25	18.9	18.0	12.2	16.9	11.2	21.1	20.2	11.9	13.3	20.9	21.2	8.8
50	23.3	20.3	14.2	23.2	20.6	20.5	25.0	16.7	20.1	23.7	28.1	18.4
100	24.7	20.0	13.8	24.1	22.6	17.8	26.2	17.6	21.4	24.3	31.4	25.1
250	25.2	21.6	15.0	25.1	24.9	16.0	26.6	17.6	26.2	27.4	34.7	33.5
500	27.9	26.2	17.4	32.3	31.9	17.2	25.7	21.8	28.2	25.3	36.6	36.4
1000	27.7	30.2	17.2	35.4	34.8	17.7	31.0	23.0	30.0	26.3	39.6	39.1
AR case: $\rho = -0.5$												
25	19.3	19.5	13.1	18.1	14.3	18.7	25.0	15.6	16.3	26.6	27.5	13.7
50	23.1	21.4	13.7	25.4	21.5	17.8	27.1	20.4	21.9	21.7	28.6	24.1
100	24.2	19.3	15.0	25.0	24.0	16.5	27.2	19.9	22.8	23.5	31.4	29.8
250	25.9	21.7	16.4	23.5	24.3	15.1	28.8	20.5	27.2	28.0	36.4	35.6
500	27.8	26.2	18.1	31.9	30.9	16.3	27.8	21.6	28.2	26.0	37.6	38.1
1000	27.4	29.2	16.6	34.6	34.1	17.1	29.2	22.9	30.3	26.5	38.6	39.0
AR case: $\rho = 0.5$												
25	7.6	8.4	4.0	13.0	1.9	13.6	12.0	14.6	11.0	23.7	0.6	11.7
50	13.6	14.4	6.2	15.7	11.5	13.7	11.8	11.2	4.8	9.5	3.0	14.0
100	20.0	19.3	11.7	22.8	21.1	14.2	22.2	21.2	14.7	22.1	20.5	13.0
250	24.6	21.9	14.9	24.8	24.1	14.5	30.7	26.8	26.7	34.1	33.4	14.4
500	26.5	25.5	17.2	32.6	32.0	15.5	29.9	27.4	25.9	35.7	35.6	14.6
1000	26.9	29.3	16.4	34.1	33.4	16.5	30.7	29.0	26.1	38.3	39.2	13.5
AR case: $\rho = 0.8$												
25	1.0	1.2	3.3	4.1	0.8	12.3	0.2	0.3	2.1	3.3	0.2	11.4
50	8.4	9.7	6.6	15.1	5.4	12.1	4.2	5.9	4.6	6.7	1.2	11.0
100	14.3	14.4	10.2	17.4	13.6	12.7	14.4	14.8	11.1	17.4	12.0	11.1
250	22.1	20.4	13.6	23.6	23.2	13.6	23.9	22.9	19.6	27.0	25.7	11.9
500	24.5	24.6	15.5	30.4	29.5	14.6	26.4	26.1	21.7	29.9	29.5	13.6
1000	26.8	30.3	15.6	35.6	34.4	16.5	27.9	29.3	24.8	36.9	36.8	13.2

Notes: See Table 3.4.

3.4 Empirical illustration: PPP of G7 economies

3.4.1 Theoretical background

The purchasing power parity (PPP) hypothesis is an important building block of many international macroeconomic models. Strong form PPP postulates that one basket of goods has the same price across different countries when expressed in a common currency. The real exchange rate of country i at time t , Q_{it} , is defined as

$$Q_{it} = \frac{E_{it}P_{it}}{P_t^*}, \quad (3.12)$$

3.4 Empirical illustration: PPP of G7 economies

where E_{it} denotes the nominal exchange rate and P_{it} and P_t^* are the domestic and foreign (numeraire) price levels, respectively. Under strong form PPP, one would expect that $Q_{it} = 1$ in all periods t . However, there are many practical reasons why strong form PPP is too hard an assumption (see e.g. Wagner, 2008 and the references therein). Hence, empirical analyses of PPP usually test for prevalence of weak form long run PPP which can be characterized as stationary real exchange rates, fluctuating around some constant mean value. It is common in the empirical literature on PPP to formalize the model in log terms and using the United States as numeraire country. Then, the real exchange rate of country i is given as

$$q_{it} = e_{it} + p_{it} - p_{us,t}^*, \quad (3.13)$$

with lower case letters denoting logs of the variables defined in (3.12). There is a vast literature on the empirical validity of PPP (see Taylor and Taylor, 2004 for a survey). Most studies based on univariate unit root tests conclude that PPP does not hold. Since it has been argued that the inability of rejecting the unit root null hypothesis might be due to low power of univariate unit root tests, panel techniques have been employed which generally yield results more in favor of long run PPP. More recently however, it has been repeatedly pointed out that results in favor of PPP obtained via first generation panel unit root tests might be plagued by cross sectional dependence and that evidence based on appropriate second generation panel unit root tests yield less support for long run PPP to hold (e.g. Wagner, 2008).

It is acknowledged that a thorough empirical investigation of PPP should make use of (second generation) panel unit root tests. However, the empirical application herein is used to highlight differences of the analyzed test procedures in an empirical context.

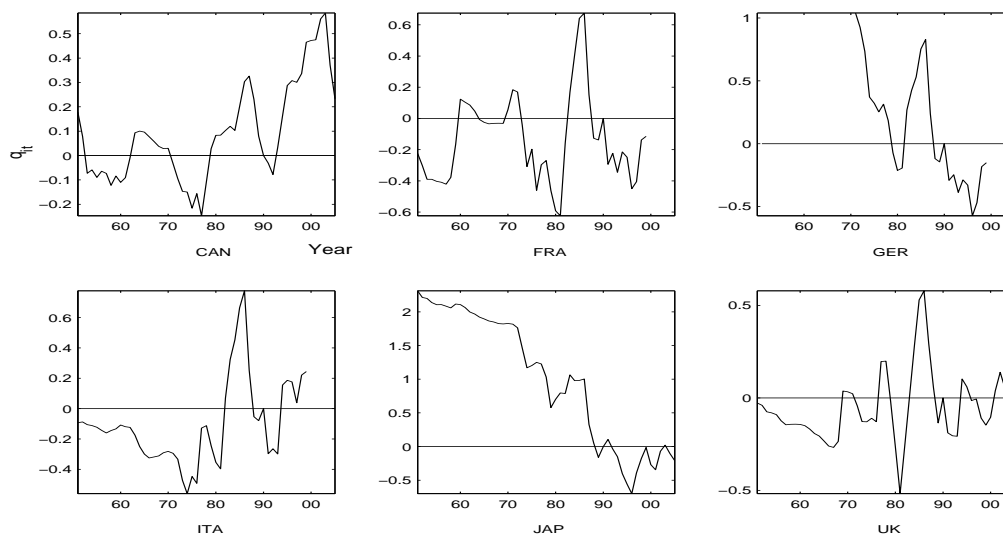
3.4.2 Data

Annual data on price levels and exchange rates are obtained from the Penn World Tables (PWT), Mark 6.2. Data spans from 1950-2004 and the base year is 2000. However, for those economies adopting the Euro in 1999, the sample is restricted to the observations prior to the introduction of the Euro. In the case of Germany, price level data availability only starts in 1970. Figure 3.2 displays the log real

3.4 Empirical illustration: PPP of G7 economies

exchange rates of the G7 economies where the United States are used as the numeraire.

Figure 3.2: Real exchange rates of G7 economies



Notes: Figures display the logs of bilateral real exchange rates versus the US-\$.

All graphs are normalized to the base year 1990.

3.4.3 Results

Construction of the test statistics proceeds as outlined in Section 3.3. According to the economic theory, only a constant term is included in the test regressions. For all three simulation based test statistics $J_{0.1}^{Reg}$, $J_{0.1}^{GLS}$ and $J_2(4)$, the mean value of 500 replications is reported in order to reduce the impact of the randomness invoked through the computer generated random walks. Marginal significance levels are computed using 100000 replications of the statistics under H_0 , and the respective time series length are matched for each of sample economies. Table 3.6 lists the empirical results. Regarding the PPP hypothesis, the overall picture is in line with most of the empirical literature on PPP in that the unit root hypothesis for the real exchange rates cannot be rejected for most of the sample economies. Notable exceptions are the United Kingdom and, to a lesser extend, France. In the case of the UK, four out of six statistics obtain significant rejections of H_0

at the 1% level ($J_{0.1}^{Reg}$, $J_{0.1}^{GLS}$, ADF^G and \bar{M}^G) while a significant rejection of the unit root hypothesis at the five (ten) percent level is obtained via $J_2(4)$ (ADF). In the case of France, rejections at the 5% level are obtained via $J_{0.1}^{Reg}$ and $J_{0.1}^{GLS}$ while ADF^G , \bar{M}^G and $J_2(4)$ indicate significance at the 10% level. However, p -values for the latter three statistics are rather close to 5%. Finally, rejections of H_0 at the 10% level are obtained for Italy via $J_{0.1}^{GLS}$ and \bar{M}^G . Rejections of H_0 can be interpreted as indicating fully integrated goods markets between the respective country and the US. In this sense, it is rather surprising that it is not possible to reject the unit root hypothesis for the case of Canada.

The empirical exercise mirrors some of the Monte Carlo results. For instance, in each of the cases where a rejection of H_0 is obtained via the ADF-GLS, a substantially larger p -value is reported for the standard ADF statistic. In line with the Monte Carlo results, this fact can be interpreted as reflecting the comparatively lower power of the latter compared with the former test. Similarly, p -values obtained for ADF^G and \bar{M}^G are usually only separated by small margins, which corresponds to the respective power estimates in Section 3.3. Finally, it can be noted that the proposed simulation based testing principle yields results which are mainly in line with more established approaches, reconfirming the promising Monte Carlo results. Among the two alternative $J_{0.1}$ statistics, somewhat lower p -values are obtained for $J_{0.1}^{GLS}$ for most of the economies in question. This finding is mirrored in the slightly better power features of $J_{0.1}^{GLS}$ over $J_{0.1}^{Reg}$ in scenarios with an intercept term only. However, it has to be highlighted that Monte Carlo studies are only illustrative since it is impossible to replicate all features of the unknown DGPs of real world economic data in an experimental setup and hence, the documented differences among the various tests in the empirical illustration might remain eventually unexplained.

3.5 Conclusions

In this chapter, a new approach to unit root testing is introduced. The underlying principle for the new test is that the regression coefficient from a spurious regression (i.e. a regression of two uncorrelated random walks) has a nondegenerated limiting distribution. In contrast, if the dependent variable is stationary, the regression coefficient converges to zero in probability. To construct a feasible

Table 3.6: Empirical results

Country	T	k	$J_{0.1}^{Reg}$	$J_{0.1}^{GLS}$	ADF	ADF^G	\bar{M}^G	$J_2(4)$
CAN	55	10	2.81 (.893)	1.99 (.750)	0.31 (.977)	-0.60 (.694)	-0.95 (.696)	0.026 (.462)
FRA	49	0	0.54** (.031)	0.39** (.020)	-2.28 (.182)	-2.27* (.052)	-8.90* (.058)	0.006* (.065)
GER	29	0	1.60 (.521)	1.47 (.634)	-2.18 (.215)	-1.34 (.436)	-2.19 (.576)	0.055 (.483)
ITA	49	0	0.81 (.131)	0.57* (.075)	-1.91 (.326)	-1.95 (.106)	-7.51* (.094)	0.013 (.196)
JAP	55	1	2.66 (.870)	2.40 (.852)	-0.83 (.803)	-0.05 (.864)	-0.03 (.835)	0.122 (.939)
UK	55	0	0.40*** (.008)	0.26*** (.003)	-2.72* (.078)	-2.75** (.014)	-13.09** (.011)	0.004** (.042)

Notes: Values below T and k refer to the available time series dimension and the chosen lag length, respectively. Numbers in parenthesis are p -values. For further notes see Table 3.2.

test statistic, it is proposed to run a sequence of regressions of the appropriately scaled data on simulated random walks with Gaussian innovations. Test statistics can then be obtained as some inter quantile ranges of the resulting empirical distribution. These statistics have an invariant limiting distribution under the null hypothesis, while they converge to zero at the rate T under the alternative hypothesis. Variants of these statistics based on the range between the 5 and 95 percentile of the simulated distribution are implemented. To account for higher order serial correlation, the data is scaled by the square root of the autoregressive spectral density variance estimator proposed by Perron and Ng (1998).

By means of a Monte Carlo study, the finite sample properties of the new test are assessed. It turns out that it has favorable size properties for most of the considered data generating processes, especially for relatively small time dimensions. In contrast to standard ADF tests, removal of deterministic terms by means of GLS detrending does not substantially improve finite sample power features of the test. In terms of size adjusted local power, the proposed test is more powerful than the standard ADF test in the intercept only case, while it is slightly less powerful than the ADF-GLS test of Elliott et al. (1996) and the \bar{M}^{GLS} test of Ng and Perron (2001) in large samples. However there are some scenarios of small samples with residual autocorrelation in which the proposed test yields highest power among those tests which achieve reasonable rejection

frequencies under H_0 . In an empirical illustration on PPP among G6 economies, it is shown that the proposed test tends to yield similar results as the most powerful benchmark tests.

A number of interesting issues are open for future research. First and foremost, the analytical derivation of the proposed test's limiting distribution deserves further consideration. Furthermore, it is not clear if the analyzed statistics are the most efficient implementation of the proposed testing principle. One could, for instance, consider alternative regression designs, e.g. implementing a regression on a multivariate random walk, use different inter quantile ranges to construct test statistics or apply other variance estimators to cope with residual serial correlation. Moreover, it should be straightforward to apply the proposed testing idea to the fields of stationarity and cointegration testing as well as to expand it to the panel case. Especially the latter appears promising, considering the relatively good performance of the proposed test in small samples. Another important issue for further research is to analyze in how far the new approach copes with violations of standard assumptions, such as outliers, breaks in the intercept or trend function as well as nonstationary volatility.

Chapter 4

A Review of Homogenous Panel Unit Root Tests

4.1 Introduction

The increasing availability and growing use of macroeconomic panel data has spurred a huge amount of research on panel unit root tests (PURT) since the early 1990ies. In typical macroeconomic applications with annual data, available time series are rather short, leading to low power of univariate unit root tests, especially under nearly integrated alternatives. Making use of cross sectional information by pooling the data or averaging over individual statistics yields tests with significantly improved power features compared with univariate tests. By now, PURTs are a standard tool in many areas of applied econometrics, but especially so in macroeconomics. The importance of PURTs in empirical macroeconomics is due to the fact that various macroeconomic hypotheses make specific assertions about the degree of integration of some key economic variables. Probably the most prominent example is the purchasing power parity hypothesis, postulating stationarity of real exchange rates. Further examples include the Fisher hypothesis which implies stationarity of real interest rates, the permanent income hypothesis stating random walk behavior of private consumption or the hypothesis of growth convergence which postulates converging (i.e. stationary) differentials of GDP growth across economies.

4.2 The autoregressive panel model

However, making use of the cross sectional dimension brings about its own set of specification issues. For instance, early (so-called first generation) tests are derived under the assumption of independent cross sectional units. As it turns out, this assumption is not only overly restrictive in most applied work, but its violation leads also to severe size distortions of first generation tests. Other PURT specification issues are related to the treatment of heterogenous intercept or trend parameters and the removal of residual serial correlation. This chapter gives a concise review over the field of PURTs, thereby providing the theoretical background for the following chapters. Thus, a particular focus is put on PURTs based on a pooled Dickey-Fuller (DF) regression, usually referred to as *homogenous* PURTs. Various first and second generation tests which are omitted in the following exposition are discussed at length in survey articles of Baltagi and Kao (2000), Hurlin and Mignon (2007) or Breitung and Pesaran (2008).

4.2 The autoregressive panel model

Recall the AR(1) time series model (2.1) introduced in Chapter 2. The panel extension of this model is given by

$$y_{it} = d_{it} + x_{it}, \quad x_{it} = \rho_i L x_{it} + u_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, N, \quad (4.1)$$

in which $d_{it} = \mu_i' z_{it}$ and the index i denotes the cross sectional unit. As before, x_{it} is the autoregressive, stochastic part of the process. The deterministic components are collected in d_{it} , where z_{it} is a $p \times 1$ vector of constants and time polynomials and μ_i is the corresponding individual specific parameter vector. Almost all PURT statistics are constructed against the homogenous null hypothesis defined by $H_0 : \rho_i = \rho = 1$,¹ whereas the tests differ with respect to the considered alternative hypotheses. In particular, two conceptually different hypotheses can be distinguished, namely

$$\begin{aligned} H_{1a} : \quad & \rho_i = \rho < 1, \\ H_{1b} : \quad & \rho_i < 1, \quad i = 1, \dots, N - R, \quad \rho_i = 1, \quad i = N - R + 1, \dots, N. \end{aligned}$$

¹A notable exception is the S_{max} statistic proposed by Chang and Song (2002).

4.2 The autoregressive panel model

While alternative H_{1a} describes a fully stationary panel with a homogenous rate of mean reversion for all cross sectional units, H_{1b} defines a heterogenous alternative under which only a fraction of the cross sectional units displays mean reverting behavior (with potentially different adjustment speeds). Accordingly, tests constructed against H_{1a} are usually called *homogenous* PURTs whereas tests constructed against H_{1b} are referred to as *heterogenous*. Homogenous PURTs are designed around a pooled DF regression (e.g. Levin et al., 2002 (LLC henceforth)²), whereas heterogenous tests are constructed by averaging over individual DF t -ratios (e.g. Im et al., 2003, (IPS in the following) Pesaran, 2007) or by combining individual specific p -values by means of Fisher (1954) type tests (for instance, Maddala and Wu, 1999 or Choi, 2001, 2006). While it has been argued that heterogenous PURTs are favorable because they are based on less restrictive assumption about the alternative hypothesis (see e.g. Maddala and Wu, 1999), homogenous tests are also consistent under heterogenous alternatives. In fact, it is shown in Breitung and Westerlund (2009) that the local power of both the homogenous LLC test and the heterogenous IPS PURT depends solely on the mean value of the individual specific autoregressive coefficients. Hence, heterogenous PURTs do not necessarily exploit the degree of heterogeneity under H_1 . In fact, Breitung and Westerlund (2009) prove that the pooled tests are generally preferable in terms of local power compared with the IPS type tests. The fact that even homogenous PURTs have power against the heterogenous alternative H_{1b} raises some ambiguity regarding the interpretation of a rejection of H_0 . Since H_{1b} allows some (non-zero) fraction of the cross sectional to display unit root processes, one cannot interpret a rejection of H_0 as indicating stationarity of the process for all cross sectional members. As Breitung and Pesaran (2008) point out, one can at most conclude that “a significant fraction of the cross section units is stationary”.

²The proposed statistic was already available in the literature from a working paper version, dating back to 1992.

4.3 First generation homogenous PURTs

4.3.1 The basic model

Consider the simplest case of model (4.1), in which $d_{it} = 0$ for all i and t . Following Breitung and Westerlund (2009), the exposition can be facilitated by introducing the following strong assumptions about the error term u_{it} ,

Assumption 4.1 \mathcal{A}_1

- (i) $E(u_{it}) = 0$ for all i and t ;
- (ii) $E(u_{it}u_{jt}) = 0$ for all i, j and t ;
- (iii) $E(u_{it}u_{is}) = 0$ for all i, t and s ;
- (iv) $E(u_{it}^2) = \sigma_u^2$ for all i and t ;
- (v) $E[e_{it}e_{jt}e_{kt}e_{lt}] < \infty$ for all i, j, k, l ,

While the mean zero assumption in $\mathcal{A}_1(i)$ is a common requirement for error terms, the absence of cross sectional correlation stated in $\mathcal{A}_1(ii)$ is the defining characteristic of first generation PURTs. The assumption of serially independent error terms in $\mathcal{A}_1(iii)$ is relaxed in Section 4.3.3. Assumption $\mathcal{A}_1(iv)$ imposes the strong assumption of cross sectional homoskedasticity which is subsequently relaxed and moreover rules out time varying variance patterns which are in the focus of Chapters 6 and 7. Finally, $\mathcal{A}_1(v)$ imposes finite fourth order moments, required for weak convergence of the test statistics.

Under model (4.1) with $d_{it} = 0$ and assumptions $\mathcal{A}_1(i)$ - $\mathcal{A}_1(v)$, H_0 can be tested by means of the pooled DF statistic obtained from the OLS regression of

$$\Delta \mathbf{y}_t = \phi \mathbf{y}_{t-1} + \mathbf{u}_t, \quad (4.2)$$

with $\phi = \rho - 1$ and $\Delta \mathbf{y}_t = (y_{1t} - y_{1,t-1}, \dots, y_{Nt} - y_{N,t-1})'$, $\mathbf{y}_t = (y_{1,t}, \dots, y_{N,t})'$ and $\mathbf{u}_t = (u_{1t}, \dots, u_{Nt})'$ are $N \times 1$ vectors. The resulting pooled DF PURT statistic is given by

$$t_{OLS} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\hat{\sigma}_u \sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{y}_{t-1}}}, \quad (4.3)$$

with $\hat{\sigma}_u^2 = (NT)^{-1} \sum_{t=1}^T (\Delta \mathbf{y}_t - \hat{\phi} \mathbf{y}_{t-1})' (\Delta \mathbf{y}_t - \hat{\phi} \mathbf{y}_{t-1})$. Under the assumed cross sectional homoskedasticity, it can be shown (see e.g. Baltagi and Kao, 2000)

4.3 First generation homogenous PURTs

that t_{OLS} converges weakly to a Gaussian limiting distribution as T and $N \rightarrow \infty$. In the fully asymptotic case with $N \rightarrow \infty$, t_{OLS} is actually a sum of N suitably scaled random variables with mean zero and unit variance. Asymptotic normality follows by application of the Lindeberg-Levy central limit theorem. All statistics reviewed in the following can be interpreted as generalizations of this simple pooled OLS statistic and share the same Gaussian limiting distribution. Therefore, the limiting distribution of the remaining statistics presented in this chapter are not explicitly stated.

Under the restrictive assumptions $\mathcal{A}_1(ii)$ and $\mathcal{A}_1(iii)$, the test statistic proposed by LLC reduces to (4.3). However, the statistic of LLC has been proposed for more general cases, including deterministic terms and residual serial correlation as well as cross sectional heteroskedasticity. The latter case is defined by the following assumption

Assumption 4.2 $\widetilde{\mathcal{A}}_1(iv)$

$E(u_{it}^2) = \sigma_{ui}^2$ for all i and t .

LLC account for cross sectional heteroskedasticity by standardizing the data with estimates of the cross section specific standard deviations,

$$t_{LLC} = \frac{\sum_{t=1}^T \tilde{\mathbf{y}}'_{t-1} \Delta \tilde{\mathbf{y}}_t}{\sqrt{\sum_{t=1}^T \tilde{\mathbf{y}}'_{t-1} \tilde{\mathbf{y}}_{t-1}}}, \quad (4.4)$$

with $\tilde{\mathbf{y}}_t = (y_{1t}/\hat{\sigma}_{u1}, \dots, y_{Nt}/\hat{\sigma}_{uN})'$, $\Delta \tilde{\mathbf{y}}_t = (\Delta y_{1t}/\hat{\sigma}_{u1}, \dots, \Delta y_{Nt}/\hat{\sigma}_{uN})'$ and $\hat{\sigma}_{ui}^2 = (T-1)^{-1} \sum_{t=1}^T \hat{u}_{it}^2$.

4.3.2 Deterministic Terms

The treatment of individual specific (or incidental) deterministic terms invokes a substantial difference in the construction of the PURT statistics compared with the univariate case. Consider model (4.1) with fixed effects or time trends, defined by $z_{it} = 1$, respectively $z_{it} = (1, t)'$. LLC propose to extend the DF approach of least squares demeaning, respectively detrending, to the panel case. However, least squares estimation of individual specific intercept or trend terms invokes the so-called Nickell bias (Nickell, 1981) which has to be accounted for by bias correction terms in the test statistic. Moreover, LLC document a substantial loss

4.3 First generation homogenous PURTs

of power for the bias adjusted t_{LLC} statistic in models with incidental intercepts or trends, compared with the case of no deterministic terms.

Breitung and Meyer (1994) and Breitung (2000) propose “unbiased” test statistics that do not require bias adjustment terms. In the case of fixed effects only (i.e. $z_{it} = 1$), Breitung and Meyer (1994) show that the incidental intercepts can be fully removed by subtracting the initial observation from the data. Thus, the statistic is given as

$$t_{BM} = \frac{\sum_{t=1}^T (\tilde{\mathbf{y}}_{t-1} - \tilde{\mathbf{y}}_0)' \Delta \tilde{\mathbf{y}}_t}{\sqrt{\sum_{t=1}^T (\tilde{\mathbf{y}}_{t-1} - \tilde{\mathbf{y}}_0)' (\tilde{\mathbf{y}}_{t-1} - \tilde{\mathbf{y}}_0)}}, \quad (4.5)$$

with $\tilde{\mathbf{y}}_{t-1}$ and $\Delta \tilde{\mathbf{y}}_t$ defined as above. It can be shown that this procedure not only circumvents the calculation of bias adjustment terms but also avoids the loss of power of the LLC statistic compared with the case of no deterministic components.

For models with incidental trends, t_{BM} is inconsistent. In order to retain asymptotic Gaussianity of the pooled test statistic without bias correction terms, Breitung (2000) suggests a detrending procedure which successfully eliminates the bias present in the least squares estimation of the trend parameters. In particular, the Helmert transformation suggested as an efficient means to center the first differences of the data in a forward looking manner, i.e.

$$\Delta y_{it}^* = s_t \left[\Delta y_{it} - \frac{1}{T-t} (\Delta y_{i,t+1} + \dots + \Delta y_{iT}) \right], \quad (4.6)$$

with $s_t^2 = (T-t)/(T-t+1)$. Detrending of the the test regression’s right hand side variable proceeds as

$$y_{it}^* = y_{it} - y_{i0} - \hat{\beta}_i t = y_{it} - y_{i0} - \frac{y_{iT} - y_{i0}}{T} t, \quad (4.7)$$

where y_{i0} and $T^{-1} \sum_{t=1}^T \Delta y_{it} = T^{-1} (y_{iT} - y_{i0})$ are used as estimators of the constants and trends, respectively. The so-called unbiased PURT statistic is then constructed from the standardized detrended data as

$$t_{UB} = \frac{\sum_{t=1}^T \tilde{\mathbf{y}}_{t-1}'^* \Delta \tilde{\mathbf{y}}_t^*}{\sqrt{\sum_{t=1}^T \tilde{\mathbf{y}}_{t-1}'^* \tilde{\mathbf{y}}_{t-1}^*}}. \quad (4.8)$$

Even though simulation results in Breitung (2000) suggest that t_{UB} is more powerful than t_{LLC} in finite samples, the introduction of individual specific trends

leads to a marked loss of power. Moon et al. (2006) point out that in a model with incidental trends, t_{UB} only has local power defined in neighborhoods shrinking at the rate of $N^{-1/4}T^{-1}$, compared with the faster rate of $N^{-1/2}T^{-1}$ in the incidental intercept case.

4.3.3 Higher order serial correlation

It is well known from the univariate case that the empirical size of unit root tests grossly deviates from the nominal level if residual serial correlation is not properly accounted for. Serial dependence is allowed for by replacing assumption $\mathcal{A}_1(iii)$ by defining u_{it} for instance as a stationary and invertible p -th order autoregressive process

Assumption 4.3 $\widetilde{\mathcal{A}}_1(iii)$

$$u_{it} = \sum_{j=1}^p \theta_{ij} u_{i,t-j} + e_{it}, \quad e_{it} \sim iid(0, \sigma_{ei}^2).$$

LLC propose the lag augmentation technique known from the univariate ADF approach in Chapter 2.3.2. After determining individual specific lag lengths, the auxiliary regressions which are run to remove the effects of deterministic components are then augmented by the respective number of lagged differences. However, if there are deterministic terms in the auxiliary regressions, this approach does not fully remove the effects of the short run dynamics from the mean of the t_{LLC} statistic. In order to obtain a Gaussian limiting distribution, the procedure of LLC requires the estimation of the ratio of the long run to short run standard deviation which enters the adjusted test statistic as a bias correction term.

Breitung and Das (2005) prove that the pooled test statistic retains its Gaussian limiting distribution without the need of bias correction terms if it is based on prewhitened data. Prewhitening proceeds by running individual specific, ADF regressions under H_0 , i.e.

$$\Delta y_{it} = \sum_{j=1}^{p_i} c_{ij} \Delta y_{i,t-j} + e_{it}. \quad (4.9)$$

The estimates $\widehat{c}_{i1}, \dots, \widehat{c}_{ip}$ are then used to obtain prewhitened data as

$$y_{it}^w = y_{it} - \widehat{c}_{i1} y_{i,t-1} - \dots - \widehat{c}_{ip} y_{i,t-p_i} \quad (4.10)$$

4.4 Second generation homogenous PURTs

and

$$\Delta y_{it}^w = \Delta y_{it} - \widehat{c}_{i1} \Delta y_{i,t-1} - \dots - \widehat{c}_{ip_i} \Delta y_{i,t-p_i}. \quad (4.11)$$

The choice of lag lengths p_i can be based on any consistent lag-length selection criterion. If the the data generating process (DGP) features both, short run dynamics and deterministic patterns, the data is first prewhitened and subsequently detrended as discussed in Section 4.3.2. Since the prewhitening regression is performed under H_0 , an intercept term has to be included only if the model includes incidental trends under the alternative hypothesis.

4.4 Second generation homogenous PURTs

Second generation PURTs are characterized by the presence of cross sectional dependence. In particular, by replacing assumption $\mathcal{A}_1(ii)$ with

Assumption 4.4 $\widetilde{\mathcal{A}}_1(ii)$

$$E(\mathbf{u}_t \mathbf{u}_t') = \Omega, \text{ with } \Omega_{ij} = \omega_{ij}, \text{ for } i \neq j \text{ and } \Omega_{ii} = \sigma_{ui}^2,$$

general forms of cross sectional correlation are permitted. Two particular forms of cross sectional dependence (i.e. static common factor and spatial autoregressive error models) are considered in more detail in the next chapter. Depending on the degree of contemporaneous co-movements, three classes of cross sectional dependence can be distinguished. Weak cross sectional dependence, strong cross sectional dependence and cross unit cointegration. The first two forms of cross sectional dependence differ with respect to the limiting behavior of the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ of Ω as $N \rightarrow \infty$. If the largest eigenvalue λ_1 is of order $O(1)$ as $N \rightarrow \infty$, Breitung and Pesaran (2008) speak of weak cross sectional dependence. Cases where λ_1 is $O(N)$ and thus diverges as N goes to infinity are classified as strong cross sectional dependence in the terminology of Breitung and Pesaran (2008). Finally, cross unit cointegration is a particular form of cross sectional dependence under which two (or more) cross sectional units form a stationary linear combination. While there are many tests available in the literature which are robust under either weak or strong form cross sectional dependence, only few tests are able to cope with cross unit cointegration (Bai and Ng, 2004 and Chang and Song, 2005).

4.4 Second generation homogenous PURTs

The literature on second generation PURTs can be divided into three directions, differing with respect to the treatment of cross sectional dependence. Firstly, many authors have developed tests by assuming a particular parametric form of cross sectional dependence in the DGP. Prominent examples of this strand of the literature are, among others, Moon and Perron (2004), Bai and Ng (2004), Choi (2006) or Pesaran (2007) who assume that the cross sectional correlation is driven by unobserved common factors. Even though factor models conceptualize strong cross sectional dependence, it has to be noted that tests which are designed under this particular form of dependence may fail to be valid under weak form dependence, as for instance, spatial error models (see e.g. Baltagi et al., 2007). Hence, other authors have suggested second generation PURTs which do not require specific parametric assumptions about the correlation pattern. Typically, such PURTs are then derived by employing generalized least squares (GLS) estimation (Harvey and Bates, 2003; O'Connell, 1998) or robust covariance estimators (Breitung and Das, 2005; Jönsson, 2005). Finally, other approaches such as instrumental variable (Chang, 2002), subsample (Choi and Chue, 2007) or bootstrap procedures (Chang, 2004; Maddala and Wu, 1999; Palm et al., 2008b) are proposed to obtain valid test procedures under unknown forms of cross sectional dependence.

4.4.1 A feasible GLS-PURT

An intuitive way of robustifying pooled PURTs against unknown forms of cross sectional dependence is to estimate the pooled DF regression by (feasible) GLS. Harvey and Bates (2003) propose FGLS estimation of the pooled DF regression in (4.2) which proceeds by premultiplying the system of equations by $\widehat{\Omega}^{-1/2}$, where

$$\widehat{\Omega} = \frac{1}{T} \sum_{t=1}^T \widehat{\mathbf{u}}_t \widehat{\mathbf{u}}_t' = \frac{1}{T} \sum_{t=1}^T (\Delta \mathbf{y}_t - \widehat{\phi} \mathbf{y}_{t-1})(\Delta \mathbf{y}_t - \widehat{\phi} \mathbf{y}_{t-1})', \quad (4.12)$$

where $\widehat{\phi}$ is the OLS estimator from 4.2. The resulting GLS PURT statistic (or multivariate homogenous DF statistic in the terminology of Harvey and Bates, 2003) is given as

$$t_{GLS} = \frac{\sum_{t=1}^T \mathbf{y}_{t-1}' \widehat{\Omega}^{-1} \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \mathbf{y}_{t-1}' \widehat{\Omega}^{-1} \mathbf{y}_{t-1}}}. \quad (4.13)$$

4.4 Second generation homogenous PURTs

The t_{GLS} statistic is asymptotically efficient and thus is asymptotically more powerful than alternative statistics which are based on the OLS estimator of ϕ . Moreover, t_{GLS} retains a Gaussian limiting distribution even under strong form cross sectional dependence as formalized by common factor models (Breitung and Das, 2008). However, a serious shortcoming of the GLS approach is that it imposes a very restrictive condition on the relative size of the time dimension, relative to the cross sectional dimension. Due to the condition that $\hat{\Omega}$ has to be invertible, t_{GLS} is only feasible if $T \geq N$. Moreover, Breitung and Das (2008) show that asymptotic Gaussianity only holds under the much stricter condition that $N^2/T \rightarrow 0$ as N and $T \rightarrow \infty$. This restrictive assumption results in poor finite sample properties (see, for instance, Breitung and Das, 2005), making the test unattractive for many situations of applied research.

4.4.2 Robust covariance estimation

In view of the restricted applicability of the statistic t_{GLS} , Jönsson (2005) and Breitung and Das (2005) independently suggest robust covariance estimation as an alternative of obtaining a cross sectional dependence robust PURT. The statistic is built on the pooled OLS estimation of (4.2) and the unknown pattern of contemporaneous correlation is approximated by so-called *panel corrected standard errors* (PCSE) (Beck and Katz, 1995). The panel corrected variance estimator of $\hat{\phi}$ is given by

$$\hat{\nu}_{\hat{\phi}} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \hat{\Omega} \mathbf{y}_{t-1}}{\left(\sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{y}_{t-1} \right)^2},$$

with $\hat{\Omega}$ as defined in (4.12). The resulting test statistic is then obtained as

$$t_{rob} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \hat{\Omega} \mathbf{y}_{t-1}}}. \quad (4.14)$$

Unlike t_{GLS} , this statistic is computational feasible even if $N > T$. It is shown in Breitung and Das (2005) that t_{rob} is asymptotically pivotal under weak form dependence. Under strong form cross sectional dependence, t_{rob} is no longer pivotal and may even diverge (Breitung and Das, 2008). However, the reported finite sample performance of t_{rob} appears to be quite robust even under strong form cross sectional dependence, often even preferable to asymptotically pivotal statistics.

4.4.3 Bootstrap PURTs

As already discussed in Chapter 2.5.2, bootstrapping unit root test statistics may yield correct inference even if the limiting distribution under H_0 is unknown or dependent on nuisance parameters. In the case of panel data, this is of particular interest since unknown patterns of cross sectional dependence invalidate first generation tests. Moreover, as seen above, pivotalness of second generation tests often hinges upon specific assumptions on the functional form of cross sectional correlation which might be hard to verify in practice. Finally, bootstrap PURTs do not require large N asymptotics and may yield asymptotic refinements compared with tests relying on asymptotic critical values. Thus far, only few bootstrap PURTs have been proposed in the literature. Most notably, Maddala and Wu (1999) and Chang (2004) consider the sieve bootstrap while a block bootstrap method is suggested by Palm et al. (2008b). Unlike in the case of the wild bootstrap presented in Chapter 2.5.2.1, these techniques generate bootstrap samples by resampling with replacement. A wild bootstrap approach to testing for unit roots in panel data is presented in the next chapter and reconsidered in Chapter 7.

The essence of bootstrapping PURT statistics is to generate bootstrap samples which preserve the cross sectional correlation present in the original data. This is achieved by vector resampling of centered residuals from individual specific first step autoregressions such as (4.9) (Chang, 2004; Maddala and Wu, 1999) or obtained from individual specific DF regressions (Palm et al., 2008b). Specifically, let $\{\tilde{e}_{it}\}_{t=1}^T = \left\{ \hat{e}_{it} - \frac{1}{T} \sum_{t=1}^T \hat{e}_{it} \right\}_{t=1}^T$ denote the sequence of centered residuals. In the case of the sieve bootstrap, a sequence of T serially uncorrelated vectors of bootstrap innovations $\mathbf{e}_t^* = (e_{1t}^*, \dots, e_{Nt}^*)'$ is drawn with replacement from $\{\tilde{e}_{it}\}_{t=1}^T$. Serial dependent bootstrap errors are then subsequently recursively constructed, employing the parameter estimates from the first step regressions (4.9). The block bootstrap method circumvents this step by resampling entire blocks of residuals which accordingly retain the serial dependence of the original data. Finally, the bootstrap sample is constructed by imposing the null hypothesis as the partial sum process of the (serially dependent) bootstrap innovations. Computing the PURT statistics for R independent bootstrap samples, with R chosen sufficiently large - in practice R is often set to 500 - yields an empirical distribution ψ_T^* from which bootstrap critical values can be obtained. If the bootstrap procedures

are asymptotically valid, the bootstrap distribution ψ_T^* asymptotically equals the true (but unknown) distribution ψ_t of the considered PURT statistic.

4.5 Conclusions

This chapter provides the background over the rapidly expanding literature on PURTs. In particular, various variants of first- and second generation homogeneous PURT statistics based on a pooled DF regression are discussed. It turns out that treatment of incidental deterministic terms and residual serial correlation poses difficulties that are not present in the univariate case. Removal of deterministic terms by least squares demeaning or trending invokes biases in the limiting distribution which can be overcome by bias correction terms (Levin et al., 2002). However, this approach induces a substantial loss of power. Demeaning and detrending procedures proposed by Breitung and Meyer (1994) and Breitung (2000) lead to more powerful tests which do not require bias correction terms. Similar arguments apply for the treatment of residual serial correlation, where the prewhitening approach of Breitung and Das (2005) appears to be preferable compared with the traditional lag augmentation known from the univariate case. A huge literature on second generation PURTs has been evolving around the issue of coping with cross sectional correlation. Here, three different approaches are presented. It is argued that the GLS test statistic, notwithstanding its theoretical merits, is outperformed in many situations of practical interest by a statistic based on the pooled DF regression and a robust covariance estimator. Finally, bootstrap PURTs are briefly introduced as an alternative means of obtaining correct inference under unknown patterns of contemporaneous dependence or finite cross sectional dimensions.

Chapter 5

Panel Unit Root Tests under Cross Sectional Dependence

5.1 Introduction

As argued in Chapter 4, panel unit root tests (PURT) are a valuable tool for applied macroeconometric research. They are not only more powerful than univariate unit root tests but can often be applied directly to test economic hypotheses. Making use of the cross sectional dimension, however, raises specification issues regarding potential contemporaneous correlation among the cross sectional units. First generation PURTs are characterized by the underlying assumption of cross sectional independence. Since neglecting cross sectional dependence leads to severe size distortions of first generation PURTs, second generation tests allow for cross sectional error term correlation. Breitung and Pesaran (2008) and Hurlin and Mignon (2007) provide recent surveys on this rapidly expanding literature. Regarding the potential sources of contemporaneous cross sectional error correlation common factor models, spatial dependence and SUR type (Zellner, 1962) approaches can be distinguished.

By means of a Monte Carlo study, Baltagi et al. (2007) analyze the performance of various PURTs under spatially dependent error terms. Spatial dependence implies that contemporaneous correlation is the stronger the closer two entities are located to each other. The concept of spatial dependence is widely

used in regional and urban economics and has a long tradition in spatial econometrics. It is thus surprising that it has only quite recently been considered in panel unit root testing since O’Connell (1998) already noted that

[...] “[A]ny EC-wide shock that influences prices or exchange rates will cause these exchange rates to move together. Or [...] shocks which originate in Germany may propagate to France but not to the U.S.”

The results of Baltagi et al. (2007) show that all analyzed tests (even the considered second generation PURTs) are to some extent sensitive to spatial autocorrelation. Rejection frequencies are generally upward distorted under the null hypothesis and the magnitude of the errors in rejection probability depends positively on the strength of the spatial correlation. These findings indicate a scope to develop test procedures which are robust under general forms of cross sectional dependence, including spatial correlation.

Complementary avenues of improving the finite sample behavior of homogenous PURTs which do not rely on a particular (dynamic) structure of cross sectional dependence are evaluated in this chapter. The focus of this chapter is throughout on the impact of (neglected) cross sectional correlation. Consequently, the simplest testing problem is investigated, namely to distinguish the panel unit root against a stationary first order autoregressive (AR(1)) alternative excluding any deterministic components. First, a PURT statistic employing a nonparametric variance estimator in the spirit of White (1980) is investigated as an alternative to the robust statistic of Jönsson (2005) and Breitung and Das (2005). Secondly, it is analyzed if the finite sample features of PURTs can be improved by constructing variance estimators from modified pooled regression residuals as suggested by MacKinnon and White (1985) and Davidson and Flachaire (2001). Thirdly, noting from Herwartz and Neumann (2005) and Herwartz (2006) that the wild bootstrap is capable to immunize test statistics against nuisance parameters invoked by SUR type disturbances, a resampling scheme for homogenous PURTs is proposed. The asymptotic validity of the wild bootstrap implementation for the simple OLS test statistic introduced in Chapter 4.3.1 is proven for the case of a finite cross sectional dimension.

The empirical features of homogenous PURT variants are studied under alternative scenarios of cross sectional dependence. To preview the simulation results,

5.2 Cross sectional dependence in panel data

it turns out that the proposed modifications offer (substantial) reductions of finite sample biases under cross sectional correlation. Wild bootstrap resampling improves the empirical features of all tests under the null hypothesis. In particular, resampling from residuals implied by the null hypothesis offers a close matching of nominal and empirical rejection frequencies under the null hypothesis. Moreover, the proposed bootstrap algorithm does not lead to reduced power under the alternative hypothesis.

As an empirical illustration the order of integration of current account imbalances is tested. The data set is sampled at the annual frequency over a period of 33 years and includes 129 economies. The example mirrors power deficiencies of univariate tests in case of small time dimensions and the impact of nuisance parameters on PURTs under contemporaneous error correlation. According to robust tests the current account (CA) to GDP ratio can be considered panel stationary.

5.2 Cross sectional dependence in panel data

5.2.1 Unit root testing in the AR(1) panel model

Reconsider the model given in (4.1). The pure AR(1) panel model

$$y_{it} = \rho_i y_{it-1} + u_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, N, \quad (5.1)$$

is obtained by setting $d_{it} = 0$. The panel unit root null hypothesis is given by $\rho_i = \rho = 1$ and is tested by means of the pooled (transformed) DF regression

$$\Delta \mathbf{y}_t = \phi \mathbf{y}_{t-1} + \mathbf{u}_t. \quad (5.2)$$

The following set of assumptions are made with respect to moment features and initial conditions.

Assumption 5.1 \mathcal{A}_1

- (i) $\mathbf{u}_t \sim iid(0, \Omega)$, with $\mathbf{u}_t = (u_{1t}, \dots, u_{Nt})'$;
- (ii) Ω is a positive definite matrix with homoskedastic diagonal elements σ_u^2 ;
- (iii) $E[u_{it}u_{jt}u_{kt}u_{lt}] < \infty$ for all i, j, k, l and t ;
- (iv) The vector of initial values $\mathbf{y}_0 = (y_{10}, \dots, y_{N0})' = 0$.

5.2 Cross sectional dependence in panel data

Assumption $\mathcal{A}_1(i)$ allows for general forms of contemporaneous error correlation but rules out serial dependence. This is admittedly a somewhat restrictive assumption, however, it has been shown in the previous chapter that it is possible to separate the issues of contemporaneous correlation on the one hand and short run dynamics on the other hand. Since the focus of this chapter is on cross sectional dependence, the simplest modeling framework abstracting from higher order serial dependence is addressed. $\mathcal{A}_1(ii)$ is also rather restrictive, imposing homoskedasticity across all cross section units. As shown in the previous chapter, however, this assumption is only required to obtain asymptotic Gaussianity of the statistic t_{OLS} under cross sectional independence and can be relaxed for the robust test statistics. Assumptions $\mathcal{A}_1(iii)$ and $\mathcal{A}_1(iv)$ are largely standard in the PURT literature and impose the existence of finite fourth order moments of the error terms and define the initial conditions. In fact, $\mathcal{A}_1(iv)$ can be relaxed to the case of initial values drawn from some stationary distribution without loss of generality.

In the following, two particular forms of cross sectional dependence are reviewed which are often encountered in applied macroeconomic modeling.

5.2.2 Common factors

Numerous authors have developed PURTs which are appropriately immunized by cross sectional dependence invoked by observed or unobserved common factors (see Hurlin and Mignon (2007) for a survey). In principle, cross sectional dependence can be invoked by one or more common factors. To simplify the exposition, however, the single factor model assumed, for instance, by Pesaran (2007) is reviewed below.

Under a common factor structure, the error term u_{it} is generated according to

$$u_{it} = \gamma_i f_t + \epsilon_{it}, \quad \epsilon_{it} \sim iid(0, \sigma_\epsilon^2), \quad (5.3)$$

where f_t is an unobserved common effect independent of ϵ_{it} with $E[f_t] = 0$. Factor loadings γ_i measure the impact of the common effect on the cross sectional unit i and ϵ_{it} are idiosyncratic error components. The contemporaneous (co)variance then depends on the factor loadings γ_i and σ_ϵ^2 . Let $\mathbf{\Gamma} = (\gamma_1, \dots, \gamma_N)'$ denote the

5.3 Finite sample modifications of homogenous PURTs

$N \times 1$ vector of factor loadings stacked over the cross section. Setting $E[f_t^2] = 1$ obtains the covariance

$$E[\mathbf{u}_t \mathbf{u}_t'] = \Omega_{CF} = \mathbf{\Gamma} \mathbf{\Gamma}' + \sigma_\epsilon^2 I_N, \quad (5.4)$$

where I_N is an $N \times N$ identity matrix. Following Breitung and Pesaran (2008) Ω_{CF} formalizes *strong* cross sectional dependence, since it can be shown that its largest eigenvalue is $O(N)$ and, thus, unbounded as $N \rightarrow \infty$.

5.2.3 Spatial dependence

Spatial modeling is appealing if contemporaneous correlation is related to some measure of location or distance. Spatially autoregressive (SAR) error terms (e.g. Elhorst, 2003) obey a representation,

$$\mathbf{u}_t = (I_N - \theta W)^{-1} \boldsymbol{\epsilon}_t, \quad |\theta| < 1, \quad \boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{Nt})', \quad \boldsymbol{\epsilon}_t \sim iid(0, \sigma_\epsilon^2 I_N), \quad (5.5)$$

where θ measures the strength of dependence and W is the (time invariant) spatial weights matrix. The structure of W is unrestricted except for the main diagonal that contains zero elements by convention. It is common practice, however, to normalize column or row sums of W to unity. One particular form of W is the '*k ahead and k behind*' structure, where shocks in one entity spill over onto the next k neighbors. For the Monte Carlo exercises this particular contiguity structure is considered. Defining $B = I_N - \theta W$, the covariance matrix is

$$E[\mathbf{u}_t \mathbf{u}_t'] = \Omega_{SAR} = \sigma_\epsilon^2 (B' B)^{-1}. \quad (5.6)$$

In contrast to the common factor model, owing to $|\theta| < 1$, the eigenvalues of Ω_{SAR} are bounded. Thus, in the terminology of Breitung and Pesaran (2008), the SAR model formalizes *weak* dependence.

5.3 Finite sample modifications of homogenous PURTs

Recall the PURT statistics t_{OLS} and t_{rob} discussed in the previous chapter. Both are based on the same pooled regression (5.2) and differ only with respect to the

5.3 Finite sample modifications of homogenous PURTs

chosen variance estimator. While t_{OLS} is valid in the current setting only if Ω is diagonal, t_{rob} is asymptotically pivotal under weak form cross sectional dependence and displays smaller finite sample size distortions than t_{OLS} in the case of strong contemporaneous correlation. In the following, an alternative cross sectional dependence robust statistic is proposed that relies on a panel generalization of the White (1980) heteroskedasticity consistent variance estimator. Moreover, MacKinnon and White (1985) and Davidson and Flachaire (2001) suggest modified regression residuals as a means of improving the finite sample properties of heteroskedasticity robust covariance estimators. This approach is applied in the construction of the modified PCSE based statistic \tilde{t}_{rob} .

5.3.1 A ‘White’ correction

In view of its construction in (4.12) it is evident that $\hat{\Omega}$ is a poor approximation of Ω in cases where $N > T$, as $N(N+1)/2$ nontrivial (co)variances are estimated using only NT distinct pieces of information. Moreover, by explicit estimation of $\hat{\Omega}$, PCSE build upon a time invariant covariance structure. Adopting a suggestion in McGarvey and Walker (2003) for stationary panel models, an alternative to the PCSE estimator is given by

$$\tilde{\nu}_{\hat{\phi}} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \check{\mathbf{u}}_t \check{\mathbf{u}}'_t \mathbf{y}_{t-1}}{(\sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{y}_{t-1})^2}, \quad (5.7)$$

where $\check{\mathbf{u}}_t$ are the residuals obtained under the null hypothesis. For the panel random walk, this amounts to using the true innovations, $\check{\mathbf{u}}_t = \mathbf{u}_t = \Delta \mathbf{y}_t$. Application of $\tilde{\nu}_{\hat{\phi}}$ in the construction of the pooled PURT statistic yields

$$t_{Wh} = \frac{\hat{\phi}}{\sqrt{\tilde{\nu}_{\hat{\phi}}}} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \check{\mathbf{u}}_t \check{\mathbf{u}}'_t \mathbf{y}_{t-1}}}. \quad (5.8)$$

Proposition 1 states the limiting distribution of t_{Wh} under the additional assumption of bounded eigenvalues of Ω from Breitung and Das (2005).

Assumption 5.2 \mathcal{A}_2

The eigenvalues of Ω as defined in \mathcal{A}_1 are $\lambda_1 \geq \dots \geq \lambda_N$ with $\lambda_1 < \bar{c} < \infty$ as $N \rightarrow \infty$.

5.3 Finite sample modifications of homogenous PURTs

Proposition 5.1 *Let the panel data generating process (DGP) be given by (5.1) and \mathcal{A}_1 to \mathcal{A}_2 . If $T \rightarrow \infty$ followed by $N \rightarrow \infty$, the statistic t_{Wh} defined in (5.8) has a Gaussian limiting distribution under H_0 .*

Proof: Proposition 5.1 follows directly from the following Lemma.

Lemma 5.1 *Under the conditions of Proposition 5.1*

$$N^{-1}T^{-2} \sum_{t=1}^T \mathbf{y}'_{t-1} \Omega \mathbf{y}_{t-1} = N^{-1}T^{-2} \sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{u}_t \mathbf{u}'_t \mathbf{y}_{t-1} + o_p(1). \quad (5.9)$$

The proof of Lemma 5.1 is given in the Appendix. It is noteworthy that the variance estimator in (5.8) also allows for time varying second order moments as it originates from robustifying common significance tests against heteroskedasticity of unknown form. This issue is considered in detail in Chapter 6. Instead of using H_0 implied residuals one may also follow the ‘typical’ White correction implemented with unrestricted regression residuals. As argued in the proof of Lemma 5.1, this approach results in the same limit distribution of t_{Wh} provided that higher order moments up to order 8 exist ($E[u_{it}^8] < \infty$).

5.3.2 Improved finite sample residuals

MacKinnon and White (1985) discuss three transformations of regression residuals that reduce the finite sample bias of heteroskedasticity consistent covariance estimators in classical (i.e. $T=1$) regression models. MacKinnon and White (1985) and Davidson and Flachaire (2001) document that a particular refinement yields most accurate bias reductions. Here, it is investigated if the adaption of the latter to the the panel case and employing it in the construction of the robust PCSE based PURT statistic in (4.14) further reduces finite sample size distortions. The preferable residual transformation (HC3 in the notation of Davidson and Flachaire, 2001) can be adapted to the panel autoregression as

$$\tilde{\mathbf{u}}_t = (\tilde{u}_{1t}, \dots, \tilde{u}_{Nt})', \quad \tilde{u}_{it} = \hat{u}_{it}/(1 - h_{it}), \quad h_{it} = y_{i,t-1}(\mathbf{y}'_{i,-} \mathbf{y}_{i,-})^{-1} y_{i,t-1}, \quad (5.10)$$

where $\mathbf{y}_{i,-} = (y_{i,0}, \dots, y_{i,T-1})'$. Replacing $\hat{\mathbf{u}}_t$ in (4.12) by $\tilde{\mathbf{u}}_t$ yields the refined test statistic \tilde{t}_{rob} . Since the residual transformation does only affect the small sample properties of the statistic, \tilde{t}_{rob} remains asymptotically Gaussian distributed.

5.4 Monte Carlo study

The finite sample properties of the new PURT statistic t_{Wh} as well as the effects of applying modified residuals in the construction of the \tilde{t}_{rob} statistic are analyzed by means of a Monte Carlo simulation study. Results for the first generation PURT statistic t_{OLS} are included as a benchmark.

5.4.1 The simulation design

Data is generated according to

$$\mathbf{y}_t = (\mathbf{1}_N - \boldsymbol{\rho})\boldsymbol{\mu} + \boldsymbol{\rho}\mathbf{y}_{t-1} + \mathbf{u}_t, \quad t = 1, \dots, T, \quad i = 1, \dots, N, \quad (5.11)$$

with

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)', \quad \mathbf{1}_N = (1, \dots, 1)', \quad \boldsymbol{\rho} = (\rho_1, \dots, \rho_N)'$$

Under H_0 , $\rho_i = 1$ for all i while under the alternative hypothesis AR parameters ρ_i are restricted to be less than unity for all cross sectional units without imposing homogeneity. Specifically, ρ_i is drawn as an *iid* sample from $U(0.96, 0.99)$ in order to guard against trivial power estimates. Note that the data generating process (DGP) allows for nonzero individual intercepts under the alternative hypothesis where, following Pesaran (2007), the fixed effects are drawn as $\mu_i \sim iid U(0, 0.02)$. Under H_1 , parameter vectors $\boldsymbol{\rho}$ and $\boldsymbol{\mu}$ are drawn only once and kept constant across all replications of an experiment. As discussed in Chapter 4.3.2 and following Breitung and Meyer (1994), the test statistics are computed on transformed data where the first observation is subtracted in order to remove the effects of the incidental intercepts.

With regard to the model disturbances the following DGPs are considered:

$$\text{DGP1: } \mathbf{u}_t = \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{Nt})', \quad \boldsymbol{\epsilon}_t \sim iid N(0, I_N)$$

$$\text{DGP2: } \mathbf{u}_t = (I_N - \theta W)^{-1} \boldsymbol{\epsilon}_t,$$

$$\text{DGP3: } \mathbf{u}_t = \boldsymbol{\Gamma} f_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\Gamma} = (\gamma_1, \dots, \gamma_N)', \quad \text{with } \gamma_i \sim iid U(0.5, 1) \forall i \\ \text{and } f_t \sim iid N(0, 1) \forall t.$$

DGP1 is the benchmark case with spherical disturbances under which all tests should have rejection frequencies close to the nominal level under H_0 . It is of interest, however, to quantify potential power losses invoked by (unnecessarily)

using robust test statistics. DGP2 formalizes spatially autocorrelated disturbances. In a first set of simulations, a row normalized spatial weights matrix W ‘*k ahead and k behind*’ ($k = 1$) is used and the spatial correlation coefficient is $\theta = 0.8$. Alternative choices for k and θ are considered in a sensitivity analysis. DGP3 addresses cross sectional dependence through one common factor. Parameters $\mathbf{\Gamma}$ and f_t are drawn only once to keep the structure of cross sectional correlation constant across all replications of an experiment.

The simulations are run for every combination of $N \in [10, 50, 100]$ and $T \in [10, 30, 50, 100]$ and each experiment is replicated 5000 times. For each experiment, 51 presample observations are generated in order to ensure convergence of the processes to their unconditional means under the alternative hypothesis. Throughout, rejection frequencies obtained under H_0 are based on the 5% Gaussian quantile. Under H_1 , reported rejection probabilities refer to test specific nominal levels such that empirical type one error probabilities equal exactly 5%. One should be careful, however, in interpreting adjusted power estimates for tests characterized by (large) violations of the nominal level under H_0 . In qualitative terms, the relative performance of alternative PURTs is unchanged when conditioning the experiments on the 1% or 10% level.

5.4.2 Rejection frequencies under H_0

Rejection frequencies for alternative tests are shown in Tables 1 (DGP1) and 2 (DGP2, upper and DGP3, lower panel). Bold entries in the Tables indicate that rejection frequencies under H_0 are not covered by a 95% confidence interval $[4.4\%, 5.6\%]$ constructed as $\alpha \pm 1.96\sqrt{\alpha(1-\alpha)/5000}$, $\alpha = 0.05$.

First consider the *iid* case. Rejection frequencies under the random walk null hypothesis show some evidence of an upward distortion for t_{OLS} and t_{Wh} in case of the smallest cross sectional dimension ($N = 10$). On the other hand, both variants of t_{rob} obtain only 1.4% and 1.0% rejections if also the time dimension is small ($T = 10$). The largest upward distortions are visible for t_{OLS} with up to 7.1% rejections. Conditional on a small panel dimension ($N, T = 10$) reliable type one error frequencies are achieved by t_{Wh} . As the cross sectional dimension increases, the overall performance of the tests improves, except for the PCSE based approaches. For panels with $N \in [50, 100]$ PCSE based statistics t_{rob} and

5.4 Monte Carlo study

\tilde{t}_{rob} are severely undersized and obtain rejection frequencies close to 5% only for values of T larger than 100 (not reported). As argued above, if $N > T$, $\hat{\Omega}$ might be a poor approximation of the true covariance matrix. It is noteworthy that the ‘White corrected’ statistic does not share this deficiency.

Table 5.1: Homogenous PURTs under cross sectional independence

N	T	H_0			H_1			$H_0(\mathbf{u}_t^* = \hat{\mathbf{u}}_t \eta_t)$			$H_1(\mathbf{u}_t^* = \hat{\mathbf{u}}_t \eta_t)$							
		OLS	rob	Wh	\widetilde{rob}	Wh	\widetilde{rob}	OLS^*	rob^*	Wh^*	OLS^*	rob^*	Wh^*					
10	10	6.5	1.4	5.0	1.0	7.3	7.7	7.4	6.7	5.0	4.9	4.4	4.7	4.5	4.4	7.7	7.7	7.3
	30	7.0	4.5	6.6	4.0	11.6	12.6	12.2	12.1	5.5	5.3	5.2	5.4	5.3	5.2	11.8	12.2	11.9
	50	7.1	5.0	6.9	5.5	21.8	21.5	21.2	20.5	4.9	5.2	5.1	5.1	5.1	5.1	20.9	20.9	19.8
	100	6.2	5.5	6.1	5.6	54.9	53.7	53.4	52.8	4.6	4.5	4.3	4.5	4.6	4.5	53.3	52.9	52.6
50	10	6.2	0.0	5.9	0.0	12.4	11.6	11.6	12.9	6.1	6.2	5.8	5.8	5.8	6.0	11.7	10.7	10.7
	30	6.4	0.5	6.3	0.4	47.4	45.2	45.2	50.3	5.7	5.7	5.8	5.8	5.8	5.4	46.4	46.3	46.3
	50	5.7	1.0	5.4	1.1	85.2	83.0	83.0	83.8	4.7	4.6	4.7	4.7	4.6	4.7	84.4	82.8	82.8
	100	5.8	2.8	5.7	2.3	100.	100.	100.	100.	5.2	5.1	4.9	5.2	5.2	5.0	100.	99.9	99.9
100	10	5.7	0.0	5.0	0.0	22.1	21.4	19.7	20.4	5.4	5.6	5.0	5.1	5.2	5.1	21.3	20.5	19.2
	30	6.0	0.0	5.7	0.0	83.6	83.3	80.6	82.3	5.2	5.3	5.2	5.3	5.2	5.2	82.8	83.0	80.6
	50	5.6	0.3	5.4	0.2	99.6	99.7	99.5	99.5	5.2	5.1	5.0	5.2	5.2	5.2	99.6	99.6	99.4
	100	5.6	1.2	5.5	0.9	100.	100.	100.	100.	5.0	5.0	4.8	4.9	4.9	4.5	100.	100.	100.

Notes: The table documents empirical rejection frequencies estimated under H_0 and H_1 , respectively. Bold numbers indicate significant deviations from the nominal 5% level. OLS denotes results obtained by the pooled OLS t -ratio, rob and Wh signify test statistics with PCSE based and ‘White corrected’ variance estimators and rob refers to finite sample residual corrections. Rejection frequencies under the alternative hypothesis are adjusted by tuning the nominal level of the tests such that rejection frequencies are exactly 5% under H_0 . Right hand side panels document empirical features of bootstrap inference indicated by ‘*’. Resampling is performed alternatively by means of unrestricted ($\hat{\mathbf{u}}_t$) and restricted ($\hat{\mathbf{u}}_t = \mathbf{u}_t$) residuals. All results are based on 5000 replications. The number of bootstrap replications is set to 499.

With respect to spatially dependent errors, the results for the OLS t -statistic are similar as reported for the LLC test in Baltagi et al. (2007) and demonstrate that spatial dependence restricts the applicability of first generation PURTs. In contrast, all considered robust tests yield empirical rejection frequencies under H_0 much closer to the nominal level. Some larger upward distortions are only notable for a small cross sectional dimension, with rejection frequencies as high as 8.6% (\tilde{t}_{rob} and $T = 100$). For medium to large cross sectional dimensions, finite sample biases decrease throughout, however, the PCSE based statistics remain prone to a downward deviation as $N \geq T$. The refined statistic \tilde{t}_{rob} does not significantly improve empirical rejection probabilities in comparison with t_{rob} and both variants of t_{rob} are generally inferior to t_{Wh} . For $T = 10$, t_{Wh} yields empirical type one errors which are covered by the 95% confidence interval constructed around the nominal level regardless of the cross sectional dimension. For larger values of T , t_{Wh} displays some moderate upward size distortion which depends inversely on the size of the cross sectional dimension.

Cross sectional dependence invoked by a common factor error model (DGP3) renders the robust tests asymptotically invalid. However, while rejection frequencies obtained by the OLS t -ratio diverge under H_0 , the robust statistic t_{Wh} yields only moderately upward distorted empirical rejection frequencies of up to 7.9%. On the other hand, the PCSE based tests display empirical type one errors below 5% except for the case of $T = 100$, with values close to the nominal 5% level.

Table 5.2: Homogenous PURTs under cross sectional correlation

N	H_0			H_1			$H_0(\mathbf{u}_t^* = \tilde{\mathbf{u}}_t \eta_t)$			$H_1(\mathbf{u}_t^* = \tilde{\mathbf{u}}_t \eta_t)$		
	OLS	rob	Wh	OLS	rob	Wh	OLS^*	rob^*	Wh^*	OLS^*	rob^*	Wh^*
Spatial AR(1) model												
10	23.6	4.1	5.4	3.4	6.3	6.1	5.4	5.5	6.0	4.7	4.9	4.8
30	24.6	6.4	7.2	6.5	9.9	9.3	8.5	8.5	5.4	5.1	4.9	4.6
50	24.9	7.3	7.7	6.6	12.9	11.7	11.8	12.5	5.4	5.1	4.9	4.6
100	24.1	8.0	8.4	8.6	26.8	20.9	21.2	19.2	5.2	5.3	5.2	5.3
50	21.5	0.3	5.6	0.3	8.6	9.0	8.2	9.3	6.3	6.1	5.7	5.2
30	22.3	2.9	7.0	2.4	21.7	21.0	20.0	23.9	5.9	5.8	5.7	5.9
50	21.2	3.6	6.4	3.5	44.5	41.5	40.4	42.4	4.7	4.9	4.7	4.8
100	21.3	5.5	6.9	5.2	85.6	81.3	80.5	80.5	5.3	5.4	5.5	5.4
100	20.8	0.0	5.5	0.0	11.0	10.5	10.9	12.4	6.4	6.1	5.7	4.9
30	21.0	0.9	5.8	1.1	41.2	41.4	37.6	40.1	5.0	4.9	5.0	4.8
50	20.4	1.9	6.4	2.3	73.8	71.6	67.8	71.6	4.8	4.9	4.9	5.0
100	20.0	3.6	6.3	3.8	99.6	99.5	99.2	99.0	4.9	5.1	4.9	5.0
Static factor model												
50	30.8	0.1	5.3	0.1	9.8	9.5	8.7	10.0	6.4	6.0	5.6	5.2
30	32.5	2.3	6.9	1.8	31.1	25.1	20.2	25.0	5.9	5.2	5.4	5.3
50	33.2	3.7	7.2	3.3	54.6	38.5	34.6	40.2	4.7	5.1	4.9	4.6
100	32.0	5.4	7.2	5.0	82.5	60.2	57.1	62.2	5.2	4.9	4.8	4.9
100	41.1	0.1	5.3	0.0	11.2	11.0	9.3	11.2	5.7	5.4	5.0	4.8
30	44.2	1.2	6.1	1.4	38.1	29.8	24.7	29.5	5.0	5.0	4.7	4.8
50	45.8	3.0	7.9	2.7	61.5	42.5	34.1	42.4	5.1	5.1	4.7	5.1
100	46.9	5.2	7.9	5.0	85.8	64.1	58.6	62.8	5.2	5.1	5.0	5.1

Notes: For factor model disturbances (DGP3) results for $N = 10$ are not provided for space considerations. Under H_0 the pattern of overrejections for this DGP is almost identical to that reported for the spatial autoregression. For further notes See Table 1.

5.4.3 Rejection frequencies under H_1

In the *iid* case (medium panel of Table 5.1), similar adjusted rejection frequencies under the alternative hypothesis are obtained for all tests. For the smallest panel dimension, rejection frequencies are around 7%, increasing quite rapidly for larger values of T and N . For panel dimensions frequently encountered in macroeconomic applications, for instance $N, T = 50$, adjusted rejection probabilities between 83 to 85.2% indicate that all tests are quite powerful, noting that the autoregressive roots in the simulation are close to unity.

Under cross sectional error correlation, results for the *OLS* statistic are unreliable, given the huge distortion of rejection frequencies under H_0 . The results for the robust tests (Table 5.2, medium panel) indicate that the power of the tests decreases in comparison with the *iid* case. For the smallest panel dimension ($N, T = 10$) and spatially correlated errors, adjusted rejection frequencies are only slightly above the nominal size of 5%. Also convergence to a unit rejection rate is slower than under cross sectional independence. While under independence a ‘certain rejection’ of H_0 is obtained for a panel dimension of $N = 50$ and $T = 100$, in the spatial error model the same holds only for $N, T = 100$. Adjusted rejection rates are even lower under common factor disturbances. In this case adjusted rejection probabilities are around 60% for the largest panel dimension.

5.4.4 Sensitivity analysis

Monte Carlo results as described above are always specific to the particular employed DGPs. Therefore, the outcome of complementary experiments with modified core features of contemporaneous error correlation are briefly sketched. For all additional Monte Carlo experiments it turned out that the relative performance of the OLS based test and its robust counterparts remain qualitatively unchanged. The reported overrejections of the null hypothesis under contemporaneous correlations turn out to depend positively on the magnitude of the spatial autocorrelation parameter θ or factor loadings $\mathbf{\Gamma}$. While empirical features of alternative tests are almost invariant for the contiguity matrices with $k = 1$ or $k = 5$, empirical overrejections under H_0 increase if the idiosyncratic error components in the common factor model (DGP3) are spatially correlated.

5.5 Bootstrap inference

5.5.1 Setup of bootstrap algorithm

The preceding analysis demonstrated the implications of asymptotic nonpivotalness of the pooled OLS test statistic in cross sectionally dependent panels. PCSE based tests or the proposed ‘White-type’ modification are asymptotically pivotal only in case of *weak* cross sectional dependence and asymptotic normality relies on the fully asymptotic case of $T \rightarrow \infty$ and $N \rightarrow \infty$. Hence, noting their failure of convergence for the case of widely used factor models, it is of immediate interest to have a test procedure at hand that works in the semiasymptotic case of a finite cross section as well as in the case of strong cross sectional dependence.

Resampling methods are often used to obtain asymptotically correct critical values for test statistics if their true limiting distribution is unknown or difficult to derive analytically. Moreover, employing bootstrap critical values might offer asymptotic refinements in comparison with first order asymptotic approximations of pivotal test statistics (Horowitz, 2001). Considering the issue of testing for individual effects in stationary panel models, Herwartz (2006) shows that the wild or external bootstrap (Wu, 1986) is suitably immunized against the lack of pivotalness under cross sectional dependence. In the field of (univariate) unit root testing, sieve resampling of Dickey-Fuller type statistics has been considered by several authors (see Chapter 2.5.2), while the wild bootstrap has been considered by Cavaliere and Taylor (2008) for unit root tests under time varying volatility. As reviewed in Chapter 4.4.3, sieve and block bootstrap approaches have also been considered for panel unit root testing. In the related field of cointegration testing, Cavaliere et al. (2010) have recently introduced a wild bootstrap method coping with non-stationary volatility of underlying error distributions. In the following, the wild bootstrap is proposed as an asymptotically valid alternative for unit root testing in dependent panels with finite cross sectional dimensions. Wild resampling of homogenous PURT’s proceeds as follows:

1. Run the pooled DF regression in (5.2) by OLS and obtain residuals $\hat{\mathbf{u}}_t = \Delta \mathbf{y}_t - \hat{\phi} \mathbf{y}_{t-1}$. Calculate the corresponding PURT statistic, denoted ψ .
2. Replicate sufficiently often the following steps:

(i) draw bootstrap residuals \mathbf{u}_t^* from $\hat{\mathbf{u}}_t$ as

$$\mathbf{u}_t^* = (u_{1t}^*, u_{2t}^*, \dots, u_{Nt}^*)' = \eta_t(\hat{u}_{1t}, \hat{u}_{2t}, \dots, \hat{u}_{Nt})', \quad \eta_t \sim iid(0, 1),$$

where η_t , $t = 1, \dots, T$, is independent from the panel data;

(ii) construct the bootstrap sample \mathbf{y}_t^* according to the DGP presumed under H_0 as

$$\mathbf{y}_t^* = \mathbf{y}_{t-1}^* + \mathbf{u}_t^*, \quad \mathbf{y}_0^* = \mathbf{y}_0;$$

(iii) calculate the bootstrap version ψ^* of ψ .

3. Decision: Reject H_0 with significance α if $\psi < c_\alpha^*$, the α -quantile of ψ^* .

Numerous choices for η_t are available from the literature (Liu, 1988, Mammen, 1993). Davidson and Flachaire (2001) illustrate the particular merits of the Rademacher distribution (Liu, 1988), where

$$\eta_t = 1 \text{ with probability } 0.5 \text{ and } \eta_t = -1 \text{ with probability } 0.5.$$

As described above, wild bootstrap samples are generated from residual estimates $\hat{\mathbf{u}}_t$ implied by $\hat{\phi}$. To improve the finite sample properties of the resampling scheme, it appears sensible to generate bootstrap innovations from restricted residuals obtained under H_0 , i.e. $\mathbf{u}_t^* = \check{\mathbf{u}}_t \eta_t$, $\check{\mathbf{u}}_t = \Delta \mathbf{y}_t$. Note that under H_0 in the AR(1) model, employing $\check{\mathbf{u}}_t$ and the Rademacher η_t amounts to drawing bootstrap series $\{\mathbf{y}_t^*\}_{t=1}^T$ from the ‘true’ residuals \mathbf{u}_t . Davidson and Flachaire (2001) point out this particular merit of the Rademacher distribution for significance testing in regression models comprising only one explanatory variable. While promising improved empirical type one error probabilities, the use of $\check{\mathbf{u}}_t$ may go at the cost of power loss. Evaluating the latter is an issue of the Monte Carlo analysis.

5.5.2 Asymptotic validity

To state the asymptotic features of the wild bootstrap scheme denote $\mathcal{Y}_T = (\mathbf{y}_0, \mathbf{u}_1, \dots, \mathbf{u}_T)$. It is shown that conditional on \mathcal{Y}_T , the bootstrap counterpart t_{OLS}^* has the same asymptotic distribution as t_{OLS} . Since \mathcal{Y}_T is itself random, weak convergence of these distributions to their common limit is obtained in probability.

It is well established that the asymptotic distribution of unit root test statistics as the DF test can be traced back to an invariance principle for suitably rescaled partial sum processes. The generalization of the invariance principle towards the multivariate case is straightforward, i.e. under the assumptions in \mathcal{A}_1 one obtains

$$B_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{u}_t \xrightarrow{d} B, \quad (5.12)$$

where $[Tr]$ is the integer part of Tr and $B = (B_1, B_2, \dots, B_N)'$ is a vector Brownian motion with covariance Ω .

The following proposition states an invariance principle for the bootstrap counterpart of $B_T(r)$.

Proposition 5.2 *Let $\mathcal{Y}_T = (\mathbf{y}_0, \mathbf{u}_1, \dots, \mathbf{u}_T)$ where vector innovations \mathbf{u}_t and initialization \mathbf{y}_0 are in accordance with \mathcal{A}_1 . Moreover, $\eta_t \sim iid(0, 1)$ and independent of \mathcal{Y}_T . Then, as $T \rightarrow \infty$ and conditional on \mathcal{Y}_T*

$$B_T^*(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \eta_t \mathbf{u}_t \xrightarrow{d_p} B, \quad (5.13)$$

where ' $\xrightarrow{d_p}$ ' signifies convergence in distribution according to the bootstrap probability measure).

As set out in Proposition 5.2 the bootstrap invariance principle holds for (true) vector innovations that might be obtained under H_0 by imposing the panel unit root. Similarly, an invariance principle can be stated if resampling is conducted with estimated panel AR(1) residual vectors. Its derivation and the proof of Proposition 5.2 are given in the Appendix.

The asymptotic validity of wild bootstrap inference can be established from the invariance principle in (5.12). The asymptotic distribution of the t_{OLS} statistic in (4.3) under H_0 with $T \rightarrow \infty$ is obtained by the continuous mapping theorem (Chang, 2004)

$$t_{OLS} \xrightarrow{d} \left(\sigma^2 \sum_{i=1}^N \sum_{j=1}^N \int_0^1 B_i B_j \right)^{-1/2} \left(\sum_{i=1}^N \int_0^1 B_i dB_i \right) \equiv \Xi, \quad (5.14)$$

where $\sigma^2 = \frac{1}{N} \sum_{i=1}^N E[u_{it}^2]$. Since $\widehat{\sigma}^2 \xrightarrow{p} \sigma^2$ and owing to (5.13) it holds accordingly that, conditional on \mathcal{Y}_T ,

$$t_{OLS}^* \xrightarrow{d} \left(\sigma^2 \sum_{i=1}^N \sum_{j=1}^N \int_0^1 B_i B_j \right)^{-1/2} \left(\sum_{i=1}^N \int_0^1 B_i dB_i \right). \quad (5.15)$$

The following Proposition states the asymptotic validity of the wild bootstrap for the case of the t_{OLS} statistic.

Proposition 5.3 *Let $\mathbf{y}_t = (y_{1t}, \dots, y_{Nt})'$ be a panel AR(1) process, generated according to (5.1) and \mathcal{A}_1 . Then, under $H_0 : \rho_i = 1, \forall i$, with N fixed and as $T \rightarrow \infty$*

$$\sup_{-\infty < g < \infty} |P(t_{OLS}^* \leq g \mid \mathcal{Y}_T) - P(\Xi \leq g)| \xrightarrow{p} 0,$$

where t_{OLS}^* is the wild bootstrap variant of t_{OLS} defined in (4.3) and Ξ the limit distribution in (5.14).

Thus, it is asymptotically irrelevant if critical values for a particular test statistic are taken from the distribution in (5.14) or generated by wild bootstrap resampling. While the distribution in (5.14) heavily depends on nuisance parameters, resampling enables an analyst to determine critical values by simulation.

Basically, the presence of nuisance parameters is invoked by finiteness of the cross sectional dimension and the fact that Ω formalizes general patterns of cross sectional heteroskedasticity and contemporaneous correlation. The latter source of nonpivotalness of t_{OLS} might be overcome along the lines in Breitung and Das (2005) in the case of *weak* dependence. Subjecting the asymptotically pivotal test statistics t_{rob} and t_{Wh} to resampling may offer faster convergence to the limit distribution (Horowitz, 2001). Finiteness of N , however, seems prohibitive for deriving analytic first order approximations of the distribution of t_{OLS} .

5.5.3 Finite sample performance

In the Monte Carlo study, the number of bootstrap replications is set to 499. Simulation results for the bootstrap implementation are documented in the right

hand side panels of Table 5.1 and 5.2. Adjusted power estimates for bootstrap tests with $\mathbf{u}_t^* = \widehat{\mathbf{u}}_t \eta_t$ are qualitatively identical to those obtained from an implementation with restricted residuals $\mathbf{u}_t^* = \check{\mathbf{u}}_t \eta_t$. Hence, only the latter are documented for space considerations. Similarly, resampling results for the modified statistic \widetilde{t}_{rob}^* are not tabulated as they are quantitatively almost identical to those obtained for t_{rob}^* .

Under independence, the wild bootstrap yields faster convergence of empirical rejection frequencies to the nominal significance level than the first order asymptotic approximation. If resampling is based on estimated residuals, a total of seven significant (but rather modest) violations of the nominal 5% level can be observed while this number is further reduced to five if resampling is based on the true residuals.

For the spatial autoregressive model, rejection frequencies under the panel unit root are similar to the results obtained under cross sectional independence, and throughout close to the nominal level. Using estimated residuals yields a rejection rate of 6.3% for t_{OLS}^* in case of the smallest panel dimension ($N, T = 10$). The results for t_{Wh}^* are supportive for resampling pivotal statistics as no significant deviation from the nominal 5% level can be detected. Generating the bootstrap sample from $\mathbf{u}_t^* = \eta_t \check{\mathbf{u}}_t$ offers further improvements of empirical rejection frequencies under H_0 . The maximum rejection rate obtained under this particular resampling scheme is 5.9% (t_{rob}^* , $N = 50, T = 30$). For DGP3, application of restricted residuals is again preferable, as all empirical rejection frequencies are covered by the 95% confidence interval constructed around the nominal level.

Results obtained under H_1 indicate that t_{OLS}^* yields the highest rejection frequencies among the considered tests. Compared with the robust tests, differences in rejection frequencies are negligibly small in the *iid* case. The relative advantage becomes more pronounced in the case of cross sectionally correlated panels and, in particular, for the common factor model. In this case, the relative power advantage of t_{OLS}^* over the least powerful variant, t_{Wh}^* , is sizeable, varying between 20 and 85 percent, with an absolute maximum difference of 29.1 percentage points ($N = 100, T = 50$). Advantages over t_{rob}^* are also sizeable but not quite as pronounced as compared with t_{Wh}^* .

Comparing the performance of bootstrap tests under H_1 with their raw counterparts, it turns out that there is hardly any power loss invoked by the resampling procedure. This is in line with the theoretical result in Cavaliere and Taylor (2008), stating that the asymptotic power function of (univariate) unit root test is not affected by the wild bootstrap.

5.5.4 Summary of Monte Carlo results

Summarizing the simulation results, the importance of accounting for cross sectional dependence is reaffirmed. While under H_0 and independence all considered tests are asymptotically pivotal, finite sample size distortions are most pronounced for PCSE based statistics. If N is relatively large compared with T , rejection frequencies obtained for t_{rob} and \tilde{t}_{rob} are substantially lower than the nominal significance level. In contrast, the ‘White’ corrected statistic t_{Wh} with restricted residuals yields (almost) correct inference in these scenarios. Inference for correlated data has to be based on robust tests, where t_{Wh} is preferable to t_{rob} and \tilde{t}_{rob} if the time dimension is relatively small. The wild bootstrap further reduces empirical type one errors in finite samples. Compared with inference based on Gaussian quantiles, resampling involves substantially improved accuracy of empirical type one errors under a common factor model. Moreover, in this scenario application of t_{OLS}^* also offers a substantial advantage in terms of (adjusted) power compared with t_{Wh}^* and t_{rob}^* .

5.6 Empirical application

To illustrate the performance of alternative PURT implementations, the order of integration of the CA to GDP ratio for is tested. Intertemporal reasoning suggests that, in the long run, CA deficits have to be repaid. This, in principle, justifies to run test regressions without intercept, but does not necessarily imply level stationarity of the CA. However, many economies display CA balances with non-zero means over the observation period. Hence, it appears more appropriate to compute the tests allowing for non-zero individual specific intercept terms. This is done by subtracting the first observation from the data as outlined in Chapter 4.3.2.

5.6 Empirical application

Panel time series of CA balances are likely affected by cross sectional correlation: In a two country world, the deficit of one economy has to be matched by a surplus of the other country. In a multi country setting, one can expect CA balances of countries with extensive financial and trade linkages to display higher dependence than CA balances of more closed economies. A potential caveat of testing for a panel unit root in the CA to GDP ratio may arise from cross country cointegration of saving and investment (or likewise, exports and imports) measured as fractions of GDP. Under liberalized capital or goods markets, however, the larger the considered cross section the lower is the likelihood of stable cross country common trends. Coakley et al. (1998) report that application of univariate unit root tests obtain mixed evidence while panel tests often support CA stationarity. In the light of the preceding results, it is not clear if these conflicting findings are due to the power deficiencies of univariate tests or invoked by large type one errors of first generation PURTs applied to cross sectionally dependent data.

Panel time series for the period 1970-2002 are drawn from the 2004 WDI CD-ROM and include up to 129 countries. The analyzed variable is the CA to GDP ratio. As the length of available time series differs across countries, the sample data is divided into four overlapping subsets. The first covers 14 economies with data available for the entire period. The second subset includes 50 entities with data available for the period 1975-2002 while the third covers 94 economies with time series starting from 1980. The last sample comprises 129 countries with data for the last ten observations. Single missing values are replaced by an AR(1) forecast.

Table 5.3 displays the empirical results. The first four columns show the panel dimensions, summarize results of the chosen lag lengths as well as of economy specific ADF tests. Columns five through eight document PURT statistics along with asymptotic and bootstrap p -values in parentheses and square brackets, respectively. Individual specific lag lengths are selected according to the Akaike information criterion (AIC). The mean value of the selected lag lengths is given in column three. The mean values are rather low, ranging between 0.19 and 1.14. Individual specific ADF tests are run with intercept terms and the appropriate number of lagged first differences while PURTs are computed on (individual specific) prewhitened data, as discussed in Chapter 4.3.3.

5.6 Empirical application

In the subsample with the largest time dimension ($T = 33$), univariate ADF tests yield rejections of H_0 for one half of the sample economies. In contrast, this share shrinks to just below 18% for the subsample with $T = 10$. For the intermediate time dimensions, the share of rejections ranges around 35%. The decreasing share of rejections illustrates potential power deficiencies of univariate DF tests under small time dimensions.

Table 5.3: Empirical results

T	N	$\overline{\text{AIC}}$	ADF	OLS	rob	Wh	\widetilde{rob}
33	14	1.14	7	-3.82	-1.72	-1.81	-1.70
				(.000)	(.043)	(.035)	(.044)
				[.008]	[.038]	[.042]	[.041]
28	50	0.60	18	-5.93	-2.29	-2.80	-2.29
				(.000)	(.011)	(.003)	(.001)
				[.000]	[.000]	[.000]	[.000]
23	94	0.41	32	-11.29	-3.44	-2.97	-3.40
				(.000)	(.000)	(.002)	(.000)
				[.000]	[.000]	[.000]	[.000]
10	129	0.19	23	-10.48	-1.91	-2.20	-1.69
				(.000)	(.028)	(.014)	(.045)
				[.000]	[.000]	[.000]	[.000]

Notes: T and N denote the time and cross sectional dimension, respectively. $\overline{\text{AIC}}$ is the mean value of individual specific lag lengths according to the AIC. ADF refer to the number of rejections of individual ADF tests obtained at the 5% level. PURT statistics documented in columns 5-8 refer to the corresponding PURT statistics t_{OLS} , t_{rob} , t_{Wh} and \widetilde{t}_{rob} . Values in parentheses are Gaussian p -values while numbers in square brackets are bootstrap p -values obtained.

Turning to PURT evidence, it is immediate that there is a large numerical difference between the t -ratios obtained via the first generation OLS statistic and the applied second generation tests. For all considered subsamples, t_{OLS} yields the by far lowest t -ratio, with asymptotic p -values of 0.000. In contrast, the t -ratios obtained for the robust PURTs are substantially smaller in absolute value in all scenarios. This finding might be interpreted as confirming the presence of cross sectional correlation in the data set since large absolute values obtained for t_{OLS} may reflect the oversizing observed in the Monte Carlo study. Nevertheless, results based on robust PURTs are generally in favor of rejecting H_0 . In the case of the large T subsample, both asymptotic and bootstrap p -values allow for a rejection of H_0 at the 5% level while H_0 can generally be rejected at the 1% level

for the two intermediate subsamples. Finally, for the subsample with $T = 10$, Gaussian p -values are larger than the bootstrap counterparts. This effect is more pronounced for the t_{rob} and \tilde{t}_{rob} statistics which can be seen as mirroring the undersizing observed for these statistics in settings in which $N > T$.

Summarizing, even when accounting for presumably important cross sectional dependence, PURTs indicate level stationarity of the CA to GDP ratio for alternative and overlapping sets of economies. However, it has to be noted that a rejection of the unit root null hypothesis by means of PURTs does not imply stationarity of the CA to GDP ratio for all cross sectional units. Hence, it appears safe to interpret the results as evidence in favor of stationary CA balances for a significant fraction of the considered economies. Relating this finding to issues such as intertemporal solvency or the so-called Feldstein-Horioka puzzle (Feldstein and Horioka, 1980) remains open for future research.

5.7 Conclusions

The focus of this chapter is the performance of first and second generation homogenous PURTs under contemporaneous correlation invoked through spatial autocorrelation and common factor models.

A modified second generation test statistic is proposed by implementing a panel variant of the heteroskedasticity robust covariance estimator of White (1980). The validity of a wild bootstrap approximation of the test based on the pooled OLS t -ratio is formally established. Considering the semiasymptotic case, the results merely rely on $T \rightarrow \infty$ and hold under cross sectional dependence in its weak or strong form.

In a simulation study, it is reconfirmed that the first generation PURT based on the pooled OLS t -statistic loses control over actual significance levels under cross sectional dependence. The second generation PURT suggested by Jönsson (2005) and Breitung and Das (2005) is characterized by distorted type one error probabilities in particular scenarios where the cross sectional exceeds the time dimension of the panel. It turns out that the proposed modified test statistic is particularly useful in these scenarios, with empirical type one errors close to the nominal significance level. The Monte Carlo analysis furthermore underscores the virtue of bootstrap inference. In particular under strong form cross sectional

dependence or for rather small panel dimensions, bootstrap tests are preferable to tests relying on asymptotic critical values. Moreover, rejection frequencies under H_1 are not adversely affected by the wild bootstrap. On the contrary, under strong cross sectional dependence, there is a sizeable power gain from using the bootstrap variant of the OLS t -statistic.

As an empirical application, panel nonstationarity of the CA to GDP ratio is investigated. The results indicate that the CA to GDP ratio can be considered panel stationary, however, given the ambiguity of interpreting rejections of the panel unit root hypothesis, this result does not imply stationarity of the CA to GDP ratio for all economies included in the panel but rather indicates that across economies, the CA balance on average tends to be mean reverting.

5.8 Appendix

5.8.1 Proof of Lemma 5.1

The result of Lemma 5.1 follows if

$$\frac{1}{N} \left[\frac{1}{T} \sum_{t=1}^T (\mathbf{y}_{t-1}/\sqrt{T})' (\mathbf{u}_t \mathbf{u}_t' - \Omega) (\mathbf{y}_{t-1}/\sqrt{T}) \right] \xrightarrow{p} 0. \quad (5.16)$$

Define $\Omega = \Phi \Lambda \Phi'$, where Λ is the diagonal matrix of eigenvalues of Ω and the columns in Φ are the corresponding eigenvectors. Then $\boldsymbol{\varepsilon}_t = \Lambda^{-1/2} \Phi' \mathbf{u}_t$ and $\mathbf{z}_t = \Lambda^{-1/2} \Phi' \mathbf{y}_t$ are mutually uncorrelated disturbances and random walks, respectively. The statistic in (5.16) has the following representation:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \frac{1}{T} \sum_{t=1}^T \frac{1}{T} \lambda_i \lambda_j z_{i,t-1} z_{j,t-1} (\varepsilon_{it} \varepsilon_{jt} - \delta_{ij}) &\equiv \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \frac{1}{T} \sum_{t=1}^T \zeta_t^{(ij)} \\ &= \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \bar{\zeta}^{(ij)}, \end{aligned}$$

where

$$\zeta_t^{(ij)} = \frac{1}{T} z_{i,t-1} z_{j,t-1} (\varepsilon_{it} \varepsilon_{jt} - \delta_{ij}),$$

and δ_{ij} is the Kronecker-Delta. The cross section specific random variables $\zeta_t^{(ij)}$ are martingale difference sequences with finite variances as implied by the law of iterated expectations. For instance,

$$\begin{aligned} \text{Var} \left[\zeta_t^{(ii)} \right] &= E \left[\left(\zeta_t^{(ii)} \right)^2 \right] \\ &= E \left[\left((\varepsilon_{i,1} + \varepsilon_{i,2} + \dots + \varepsilon_{i,t-1}) / \sqrt{T} \right)^4 E \left[(\varepsilon_{i,t}^2 - 1)^2 \right] \right] < \infty. \end{aligned}$$

Therefore $\bar{\zeta}^{(ij)} = 1/T \sum_{t=1}^T \zeta_t^{(ij)} = o_p(1)$. All components of the cross sectional sum $(\sum_i \sum_j \bar{\zeta}^{(ij)})$ have mean zero and $\text{Var}[\bar{\zeta}^{(ij)}] = O(T^{-1})$. Moreover, the cross sectional sum consists of uncorrelated components such that its overall variance is a sum of N^2 nonzero variances. By assumption, the eigenvalues of Ω are bounded such that $\text{Var}[\sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \bar{\zeta}^{(ij)}] = O(N^2)O(T^{-1})$. Since $E[\bar{\zeta}^{(ij)}] = 0$, $1/N \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \bar{\zeta}^{(ij)}$ is $o_p(1)$, such that the result in (5.16) applies.

Similar reasoning for the asymptotic behavior of the White corrected variance estimator also holds if estimated panel AR(1) residuals are used. For proving

$$\frac{1}{N} \left[\frac{1}{T} \sum_{t=1}^T (\mathbf{y}_{t-1}/\sqrt{T})' (\hat{\mathbf{u}}_t \hat{\mathbf{u}}_t' - \Omega) (\mathbf{y}_{t-1}/\sqrt{T}) \right] \xrightarrow{p} 0, \quad (5.17)$$

with $\hat{\mathbf{u}}_t = \mathbf{u}_t - q\mathbf{y}_{t-1}$ and

$$q = \left(\sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{y}_{t-1} \right)^{-1} \sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{u}_t, \quad (5.18)$$

it turns out, however, that convergence in probability requires higher order moment conditions as formalized by Assumption \mathcal{A}_1 . To be precise $E[u_{it}^8] < \infty$ is required for the result in (5.17). □

5.8.2 Proof of Proposition 5.3

From (5.12) it follows that the partial sum process $B_T(r)$ is multivariate normal over any subinterval. For the bootstrap approximation the result in (5.13) follows from the application of the multivariate Lindeberg-Feller central limit theorem (e.g. Greene, 2003, p. 913). Since $\eta_t \sim iid(0, 1)$ it follows that for any $0 < r \leq 1$ as $T \rightarrow \infty$

$$\frac{\sqrt{[Tr]}}{\sqrt{T}} \frac{1}{\sqrt{[Tr]}} \sum_{t=1}^{[Tr]} (\eta_1 \mathbf{u}_1 + \eta_2 \mathbf{u}_2 + \dots + \eta_{[Tr]} \mathbf{u}_{[Tr]}) \xrightarrow{d} N(0, r\bar{\Omega}_{[Tr]}). \quad (5.19)$$

The asymptotic covariance in (5.19)

$$\bar{\Omega}_{[Tr]} = \lim_{T \rightarrow \infty} (\mathbf{u}_1 \mathbf{u}'_1 + \mathbf{u}_2 \mathbf{u}'_2 + \dots + \mathbf{u}_{[Tr]} \mathbf{u}'_{[Tr]}) / [Tr]$$

exists if

$$\lim_{T \rightarrow \infty} (T\bar{\Omega}_T)^{-1} (\mathbf{u}_t \mathbf{u}'_t) = 0, \quad \forall t. \quad (5.20)$$

The latter condition holds under the moment requirement in \mathcal{A}_1 . Moreover, owing to ‘strict exogeneity’ of η_t , $B_T^*(r)$ is independent of the sample realization. Since, $\bar{\Omega}_T \xrightarrow{p} \Omega$, $B_T(r)$ and $B_T^*(r)$ share the same limit distribution. □

5.8.3 Invariance principle for the wild bootstrap based on estimated residuals

In the case of resampling from estimated panel AR(1) residual vectors, the invariance principle is given by

$$\begin{aligned}
 B_T^*(r) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{u}_t^* = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \hat{\mathbf{u}}_t \eta_t = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{u}_t \eta_t - \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{y}_{t-1} q \eta_t \\
 &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{u}_t \eta_t + o_p(1) \xrightarrow{d} B,
 \end{aligned} \tag{5.21}$$

with q defined in (5.18).

To verify (5.21) consider the rescaled discrete sum

$$\begin{aligned}
 \sum_{t=1}^T \mathbf{y}_{t-1} q \eta_t &= q \left(\mathbf{u}_1 \sum_{t=2}^T \eta_t + \mathbf{u}_2 \sum_{t=3}^T \eta_t + \cdots + \mathbf{u}_{T-2} \sum_{t=T-1}^T \eta_t + \mathbf{u}_{T-1} \eta_T \right) \\
 &= q \zeta.
 \end{aligned}$$

Obviously, $E^*[\zeta] = 0$ and $\text{Cov}^*[\zeta] = O(T^2)$. Since $1/T\zeta = O_p^*(1)$ and $q = O_p(T^{-1})$ it follows that $\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \mathbf{y}_{t-1} q \eta_t = o_p(1)$.

□

Chapter 6

Panel Unit Root Tests under a break in the innovation variance

6.1 Introduction

Chapter 5 has highlighted the effects of neglected cross sectional dependence on first generation panel unit root tests (PURT). While first generation tests assuming cross sectional independence display substantially distorted empirical type one errors, it has been shown that more reliable inference can be obtained by application of robust second generation tests.

Second order invariance of model disturbances is an additional implicit assumption of PURTs. However, this assumption is quite restrictive, as many macroeconomic and financial variables are characterized by structural shifts in their unconditional volatility. In fact, what has become known as the ‘Great Moderation’ is a substantial decline in numerous macroeconomic key variables’ volatility across the G7 economies since the mid 1980s (see, for instance, Kim and Nelson, 1999, McConnell, 2000 and Stock and Watson, 2003). The adverse effects of variance shifts on unit root tests for single time series have been studied by, among others, Hamori and Tokihisa (1997), Kim et al. (2002), Cavaliere (2004), and Cavaliere and Taylor (2007a,b, 2008). The main findings are that the (augmented) Dickey-Fuller (Dickey and Fuller, 1979, Said and Dickey, 1984 (A)DF henceforth) and other unit root tests asymptotically depend on nuisance

parameters in the presence of permanent variance shifts. Hence, seriously distorted empirical type one errors and deceptive inference are the consequences of a violation of the implicit assumption of time invariant volatility. The magnitude of size distortions is shown to depend on specific break patterns. Generally, largest (positive) size distortions are observed for early negative and late positive shifts in the level of the process' unconditional variance. So far, only Hanck (2009b) attempts to generalize these results to the field of panel unit root testing. However, while he considers intersection tests for heterogenous panels which are constructed by combining the p -values obtained from volatility break robust univariate tests, this chapter concentrates on the class of class of homogenous PURTs based on a pooled DF regression as discussed in the previous chapter. It is shown that the second generation 'White-type' corrected PURT proposed in Chapter 5.3.1 retains a Gaussian limiting distribution even under discrete shifts of the innovation variance. In contrast, the first- and second generation PURT statistics of Levin et al. (2002) and Breitung and Das (2005), respectively, do not converge to a nuisance free limiting distribution in this case. Moreover, the local asymptotic power function of the 'White' corrected test statistic is derived. It turns out that in absence of volatility breaks, its local asymptotic power equals those of the statistic proposed by Breitung and Das (2005), while in the presence of a volatility break, the power depends on the timing and direction of the break. Deterministic terms and residual serial correlation are accounted for by the detrending and prewhitening schemes proposed in Breitung (2000) and Breitung and Das (2005), respectively. While the prewhitening scheme works well even under second order moment instability, the detrending scheme invokes serious deviations of empirical type one errors from the nominal significance level if there is a break in the innovation variance.

As an empirical application, PURT based evidence on the Fisher hypothesis in Crowder (2003) is reconsidered. Postulating a one-to-one comovement of nominal interest rates and expected rates of inflation, the Fisher hypothesis implies stationary real interest rates. The considered cross section consists of 9 developed economies for which data is sample at the quarterly frequency over the period 1961Q2-2007Q2. The data set mirrors core issues discussed in this chapter, such as shifts in unconditional volatility and cross sectional dependence.

6.2 Effects of nonstationary volatility on univariate unit root tests

The effects of nonstationary volatility on unit root tests in the univariate case have been investigated by Hamori and Tokihisa (1997), Kim et al. (2002), Cavaliere (2004) and Cavaliere and Taylor (2007a,b, 2008). For illustrative purposes, the results of Hamori and Tokihisa (1997) are briefly reviewed in the following. Hamori and Tokihisa (1997) consider the most basic example of a single upward shift in the innovation variance of an autoregressive process of order one (AR(1)) without any deterministic terms. In particular, consider the following data generating process (DGP)

$$y_t = \rho y_{t-1} + u_t, \quad t = 1, \dots, T. \quad (6.1)$$

In (6.1), the variance shift is modeled by means of the composite error term u_t , i.e.

$$u_t = e_t + \eta_t DU_t, \quad e_t \sim iid(0, \sigma_1^2), \quad \eta_t \sim iid(0, \sigma_2^2)$$

$$\text{and } DU_t = \begin{cases} 1, & \text{if } t > T_B, \text{ (} 1 < T_B < T \text{)} \\ 0, & \text{otherwise,} \end{cases}$$

where T_B denotes the timing of the variance break. Let $\delta = T_B/T$ denote the ratio of pre-break to total sample period and $W(r)$ is a standard Brownian motion defined on $r \in [0, 1]$, then, as $T \rightarrow \infty$, the asymptotic distribution of the DF t -ratio of $\hat{\rho} - 1$ is given by

$$t_{DF} \xrightarrow{d} \frac{\frac{1}{2}[\{W(1)\}^2 - 1]}{\sqrt{\int_0^1 W(r)dr^2 - \Xi \left[(1 - \delta) \int_0^\delta W(r)dr + \delta \int_\delta^1 \left(\frac{1-r}{r}\right) W(r)dr \right]}}, \quad (6.2)$$

$$\text{where } \Xi = \frac{(\sigma_2/\sigma_1)^2}{1 + (\sigma_2/\sigma_1)^2 (1 - \delta)}.$$

It is easy to verify that the nuisance parameters in the limiting distribution depend on the strength and the timing of the variance break. The standard DF case is covered by $\sigma_2 = 0$ and $\delta = 0$ or $\delta = 1$. Hamori and Tokihisa (1997) provide simulation evidence suggesting that a late positive variance shift leads to the largest (upward) bias of empirical type one errors. Kim et al. (2002) generalize the previous result to models with deterministic terms and propose a

pivotal test for the unit root null hypothesis based on prior break date estimation. In a series of papers, Cavaliere (2004) and Cavaliere and Taylor (2007a,b, 2008) extend these results in three directions. First, they allow for a wider class of volatility processes, including multiple breaks and trending volatility. Second, they extend the analysis to the class of M -type of unit root tests proposed by Perron and Ng (1996), Stock (1999) and Ng and Perron (2001). Finally, they propose alternative volatility-break robust test procedures, such as a test based on the estimated variance profile, as well as tests based on simulation or resampling methods.

6.3 PURTs under nonstationary volatility

6.3.1 The autoregressive, heteroskedastic panel model

In light of the results from the univariate case, it appears promising to study the effects of nonstationary volatility on homogenous PURTs. More specifically, the limiting distributions of alternative t -statistics obtained from pooled DF regressions are derived for a panel AR(1) model allowing for multiple and possibly heterogeneous breaks in the innovation variance as well as for weak cross sectional dependence. Weak cross sectional dependence as defined by Breitung and Pesaran (2008) is characterized by bounded eigenvalues of the covariance matrix as $N \rightarrow \infty$. This type of dependence includes, for instance, covariance matrices implied by all types of spatial panel models (Elhorst, 2003) but excludes dependence invoked through common factor models. As shown in Chapter 4.3.2 and 4.3.3, the treatment of deterministic terms and serially correlated disturbances can be separated from the derivation of the test statistics' limiting distribution. Hence, the considered heteroskedastic panel model abstracts from these features relevant for most empirical applications and is given by

$$\mathbf{y}_t = \rho \mathbf{y}_{t-1} + \mathbf{u}_t, \quad t = 1, \dots, T. \quad (6.3)$$

As in the previous chapters, $\mathbf{y}_t = (y_{1t}, \dots, y_{Nt})'$, $\mathbf{y}_{t-1} = (y_{1,t-1}, \dots, y_{N,t-1})'$ and $\mathbf{u}_t = (u_{1t}, \dots, u_{Nt})'$ are $N \times 1$ vectors and the index $i = 1, \dots, N$ indicates the cross sectional units. The autoregressive coefficient ρ satisfies either $\rho = 1$ under the unit root null hypothesis or $|\rho| < 1$ under the stationary alternative hypothesis.

6.3 PURTs under nonstationary volatility

Unless otherwise stated, the following set of assumptions regarding the vector of errors \mathbf{u}_t is supposed to hold throughout

Assumption 6.1 (\mathcal{A}_1)

- (i) *The error vector $\mathbf{u}_t \sim iid(\mathbf{0}, \Omega_t)$;*
- (ii) *Ω_t is a positive definite matrix with eigenvalues $\lambda_1 \geq \dots \geq \lambda_N$ and $\lambda_1 < \bar{c} < \infty$ for all t ;*
- (iii) *Moreover, it is assumed that $E[u_{it}u_{jt}u_{kt}u_{lt}] < \infty$ for all i, j, k, l ;*
- (iv) *Finally, the vector of initial conditions is given by $\mathbf{y}_0 = (0, \dots, 0)'$.*

The assumptions $\mathcal{A}_1(i)$ - $\mathcal{A}_1(iv)$ are basically the same as in Breitung and Das (2005) or Chapter 5.2.1 except that a time varying covariance matrix, Ω_t is explicitly allowed for. $\mathcal{A}_1(i)$ rules out higher order serial correlation which will be considered later. $\mathcal{A}_1(ii)$ restricts the pattern of cross sectional dependence to the weak type dependence while the assumed existence of finite fourth order moments of u_{it} in $\mathcal{A}_1(iii)$ is a standard assumption in the (panel) unit root literature. Zero initial conditions are defined in $\mathcal{A}_1(iv)$. Additionally, the following assumptions are made with respect to Ω_t , further defining the types of volatility breaks considered in this chapter.

Assumption 6.2 (\mathcal{A}_2)

- (i) $\Omega_t = \Omega_1$ for $t = 1, \dots, T_1$ and $\Omega_t = \Omega_2$ for $t = T + 1, \dots, T$.
- (ii) *Moreover, $T_1, T_2 \rightarrow \infty$ as $T \rightarrow \infty$ with $T_2 = T - T_1$ and $\frac{T_1}{T} \rightarrow \delta > 0$, $\frac{T_2}{T} \rightarrow 1 - \delta > 0$.*
- (iii) *Finally, $\Omega_t = \Phi_t^{1/2} \Psi \Phi_t^{1/2}$ for all t , with $\Phi_t = diag(\sigma_{1t}^2, \dots, \sigma_{Nt}^2)$.*

Assumption $\mathcal{A}_2(i)$ restricts the number of variance break points to one. This assumption is made to simplify the analytical derivations. From the proofs in the Appendix, however, it will become clear that it is straightforward to modify the analytical derivations to account for multiple breaks. $\mathcal{A}_2(ii)$ requires that the pre- and post-break samples increase as $T \rightarrow \infty$, with the subsamples being some constant fractions δ and $(1 - \delta)$ of the total sample, respectively. This assumption is important in the derivation of the limiting distribution as it ensures convergence of partial sum processes to functionals of Brownian motions in each

6.3 PURTs under nonstationary volatility

subsample. Assumption $\mathcal{A}_2(iii)$ defines the type of variance break considered. The decomposition $\Omega_t = \Phi_t^{1/2} \Psi \Phi_t^{1/2}$, where Ψ is the (time invariant) correlation matrix implied by Ω_t and $\Phi_t^{1/2}$ is a diagonal matrix of the idiosyncratic standard deviations, allows to separate the issues of cross sectional dependence and variance breaks. It further incorporates heterogeneity along the cross sectional dimension as idiosyncratic variances σ_{it}^2 and the strength of the variance breaks may differ. Obviously, $\mathcal{A}_2(iii)$ also covers the case where only a fraction of the cross sectional units feature a shift in the innovation variance.

6.3.2 Asymptotic size distortions of homogenous PURTs

Consider the AR(1) panel model defined in (6.3). Recall the OLS t -statistic introduced in Chapter 4.3.1. Based on the the pooled DF regression

$$\Delta \mathbf{y}_t = \phi \mathbf{y}_{t-1} + \mathbf{u}_t,$$

with $\Delta \mathbf{y}_t = (y_{1t} - y_{1,t-1}, \dots, y_{Nt} - y_{N,t-1})'$, the test statistic is computed as

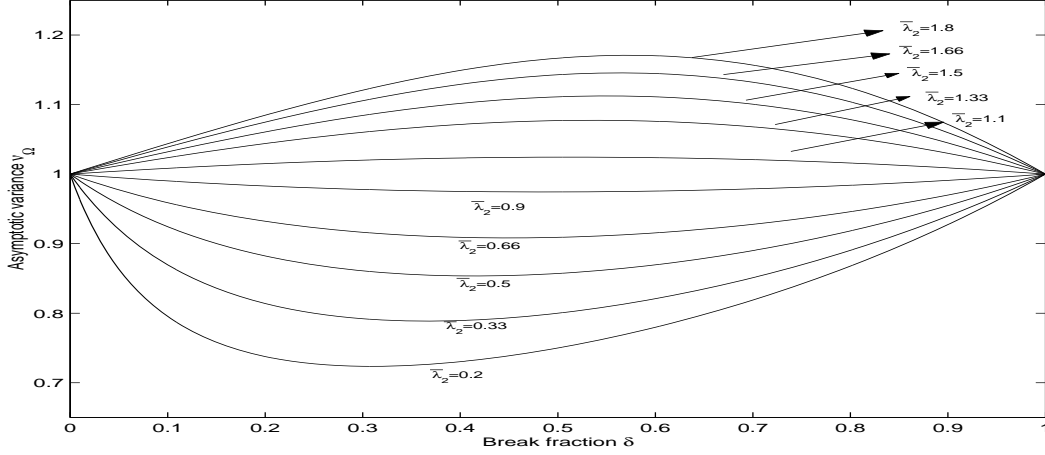
$$t_{OLS} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\hat{\sigma}_u \sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{y}_{t-1}}}, \quad \hat{\sigma}_u^2 = (NT)^{-1} \sum_{t=1}^T (\Delta \mathbf{y}_t - \hat{\phi} \mathbf{y}_{t-1})' (\Delta \mathbf{y}_t - \hat{\phi} \mathbf{y}_{t-1}) \quad (6.4)$$

Under H_0 in (6.3) with with cross sectionally independent and homoskedastic error terms with constant variance ($\Omega_t = \Omega = I_N \sigma_u^2$), the test statistic t_{OLS} is asymptotically Gaussian as $T, N \rightarrow \infty$. Violations of the assumption of cross sectional independence can be overcome along the lines outlined in Chapters 4.4.2, 5.3.1 and 5.5, either by means of robust covariance estimation or resampling methods. The effects of a break in the innovation variance on homogenous PURTs have not yet been studied. In the following, it is shown that in analogy to the univariate case, t_{OLS} does not converge to a nuisance free limiting distribution and, hence, loses control over the asymptotic size of the test.

Proposition 6.1 *Assume the panel DGP is given by (6.3) and assumptions \mathcal{A}_1 and \mathcal{A}_2 with $\Psi = I_N$ and $\sigma_{i\bullet}^2 = \sigma_{u\bullet}^2 \forall i$ and $\bullet = 1, 2$. Then, under $H_0 : \rho = 1$ and for $T \rightarrow \infty$ followed by $N \rightarrow \infty$, $t_{OLS} \xrightarrow{d} N(0, \bar{v}_{OLS})$, $\bar{v}_{OLS} \neq 1$ if $\sigma_{u1}^2 \neq \sigma_{u2}^2$.*

6.3 PURTs under nonstationary volatility

Figure 6.1: Asymptotic variance of t_{OLS} , $\bar{\nu}_{OLS}$



The proof of Proposition 6.1 is deferred to Section 6.7.1 in the Appendix. The result directly shows that discrete shifts in the innovation variance induce nuisance parameters in the asymptotic distribution of the t_{OLS} PURT statistic. Moreover, given the specific form of $\bar{\nu}_{OLS}$ derived in the Appendix, it is clear that the direction and strength of the implied size distortion depend on the specification of the break. In particular, it follows that

$$\bar{\nu}_{OLS} = \frac{0.5\delta^2\bar{\lambda}_1^2 + \delta(1-\delta)\bar{\lambda}_1\bar{\lambda}_2 + 0.5(1-\delta)^2\bar{\lambda}_2^2}{(\delta\bar{\lambda}_1 + (1-\delta)\bar{\lambda}_2)[0.5\delta^2\bar{\lambda}_1 + \delta(1-\delta)\bar{\lambda}_1 + 0.5(1-\delta)^2\bar{\lambda}_2]},$$

with $\bar{\lambda}_\bullet = N^{-1} \sum_{i=1}^N \lambda_{i\bullet}$, $\bar{\lambda}_\bullet^2 = N^{-1} \sum_{i=1}^N \lambda_{i\bullet}^2$, where $\bullet = 1, 2$ refers to the pre- and post-break period, respectively and $\bar{\lambda}_1\bar{\lambda}_2 = N^{-1} \sum_{i=1}^N \lambda_{i1}\lambda_{i2}$. To illustrate the size distortion invoked by variance breaks, Figure 6.1 depicts the asymptotic variance $\bar{\nu}_{OLS}$ for a continuity of breakpoints $\delta \in [0, 1]$ with

$$\bar{\lambda}_1 = 1 \text{ and } \bar{\lambda}_2 \in \{0.2, 0.33, 0.5, 0.66, 0.9, 1.1, 1.33, 1.5, 1.66, 1.8\}.$$

Figure 6.1 reveals that largest deviations of $\bar{\nu}_{OLS}$ from unity are characteristic for late positive and early negative variance breaks. This result is in line with findings for the time series case, where early negative and late positive variance shifts have been found to induce largest size distortions. However, in the time series case, both scenarios induce an upward size distortion, while the simulated values of $\bar{\nu}_{OLS}$ imply a downward size distortion in the case of a negative variance

6.3 PURTs under nonstationary volatility

break. This is easily seen by noting that $\bar{\nu}_{OLS} < 1$ corresponds to less probability mass in the tails of the asymptotic distribution of t_{OLS} compared with the Gaussian distribution.

So far, results are derived under cross sectional independence and homoskedasticity. However, asymptotic size distortions carry over to the cross sectional dependence robust statistic t_{rob} suggested by Breitung and Das (2005). As discussed in Chapter 4.4.2, the statistic obtained by applying *panel corrected standard errors* (Beck and Katz, 1995) is given as

$$t_{rob} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \hat{\Omega} \mathbf{y}_{t-1}}}, \quad \text{with} \quad \hat{\Omega} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}'_t. \quad (6.5)$$

Under weak form cross sectional dependence with a covariance structure characterized by bounded eigenvalues as $N \rightarrow \infty$ and time invariant innovation variance, the statistic retains a Gaussian limiting distribution. However, Proposition 6.2 states that this result does no longer hold in the case of a discrete break in the innovation variance.

Proposition 6.2 *Assume the panel DGP is given by (6.3) and assumptions \mathcal{A}_1 and \mathcal{A}_2 . Then, under $H_0 : \rho = 1$ and for $T \rightarrow \infty$ followed by $N \rightarrow \infty$, $t_{rob} \xrightarrow{d} N(0, \bar{\nu}_{rob})$, $\bar{\nu}_{rob} \neq 1$ if $\sigma_{u1}^2 \neq \sigma_{u2}^2$.*

The proof of Proposition 6.2 is given in Section 6.7.2 of the Appendix. It turns out that

$$\bar{\nu}_{rob} = \frac{0.5\delta^2 \bar{\lambda}_1^2 + \delta(1-\delta) \bar{\lambda}_1 \bar{\lambda}_2 + 0.5(1-\delta)^2 \bar{\lambda}_2^2}{(\delta^2 - 0.5\delta^3) \bar{\lambda}_1^2 + (1.5\delta + 0.5\delta^2 - 2\delta^3) \bar{\lambda}_1 \bar{\lambda}_2 + 0.5(1-\delta)^3 \bar{\lambda}_2^2}.$$

It is easy to verify that in absence of volatility breaks the result in Breitung and Das (2005) obtains as a special case with $\delta = 0$ or $\delta = 1$.

6.3.3 A volatility-break robust test

Reconsider the ‘White-type’ test statistic proposed in Chapter 5.3.1. Making use of residuals obtained under H_0 , the test statistic and its asymptotic distribution are

$$t_{Wh} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \check{\mathbf{u}}_t \check{\mathbf{u}}'_t \mathbf{y}_{t-1}}} \xrightarrow{d} N(0, 1), \quad \check{\mathbf{u}}_t = \Delta \mathbf{y}_t = \mathbf{u}_t. \quad (6.6)$$

6.3 PURTs under nonstationary volatility

Given the construction of the employed covariance estimator, one might expect that t_{Wh} is robust with respect to unknown patterns of (nonstationary) heteroskedasticity. Similarly, Hamori and Tokihisa (1997) suggest the White correction (with unrestricted residuals, however) as a potential means to appropriately cope with the nuisance invoked by a variance shift. The following Proposition states asymptotic Gaussianity of the statistic t_{Wh} under a volatility break as defined by \mathcal{A}_2 .

Proposition 6.3 *Assume the DGP is given by (6.3) and Assumptions \mathcal{A}_1 and \mathcal{A}_2 hold and $\sigma_{u1}^2 \neq \sigma_{u2}^2$. Then under $H_0 : \rho = 1$ and for $T \rightarrow \infty$ followed by $N \rightarrow \infty$, $t_{Wh} \xrightarrow{d} N(0, 1)$.*

The proof of Proposition 6.3 is derived in Section 6.7.3 in the Appendix.

Even though the proof is laid out for a single break date, it is straightforward to extend it to scenarios of multiple breaks. A caveat of the asymptotic results is that they are obtained under sequential asymptotics. As it is shown in Phillips and Moon (1999), sequential asymptotics do not necessarily imply convergence if N and T approach infinity jointly. However, results in Breitung and Westerlund (2009) conjecture that the previous results might also apply if $\sqrt{N}/T \rightarrow 0$ as $T, N \rightarrow \infty$ jointly.

6.3.4 Local asymptotic power of t_{Wh}

To verify that the test based on t_{Wh} has asymptotic power in local-to-unity neighborhoods, the following Proposition states its asymptotic distribution under a sequence of local alternatives given by

$$H_l : \rho = 1 - \frac{c}{T\sqrt{N}}. \quad (6.7)$$

Proposition 6.4 *Under the sequence of local alternatives defined in (6.7), for $T \rightarrow \infty$ followed by $N \rightarrow \infty$, t_{Wh} is asymptotically distributed as $N(-c\mu_l, 1)$, where*

$$\mu_l = \frac{0.5\delta^2\bar{\lambda}_1 + \delta(1-\delta)\bar{\lambda}_1 + 0.5(1-\delta)^2\bar{\lambda}_2}{\sqrt{0.5\delta^2\bar{\lambda}_1^2 + \delta(1-\delta)\bar{\lambda}_1\bar{\lambda}_2 + 0.5(1-\delta)^2\bar{\lambda}_2^2}}.$$

The proof of Proposition 6.4 is deferred to Section 6.7.4 in the Appendix. The result directly implies asymptotic power of the test in local-to-unity neighborhoods of order $O(T^{-1}N^{-1/2})$ for models without individual time trends. Moreover, it is easy to see that in the case of time invariant volatility with $\delta = 1$, $\mu_l = \sqrt{0.5} \bar{\lambda}_1 / \sqrt{\lambda_1^2}$, implying the same local asymptotic power as obtained by Breitung and Das (2005) for the t_{rob} statistic. Finally, a more detailed investigation of μ_l reveals that a downward (upward) shift of the innovation variance leads to asymptotically higher (lower) local power compared with the benchmark case of constant volatility.

6.4 Monte Carlo study

6.4.1 The simulation design

To illustrate the finite sample effects of volatility breaks on the considered homogenous PURTs, three stylized scenarios are considered:

$$\begin{aligned} \text{DGP1: } \mathbf{y}_t &= (1 - \rho)\boldsymbol{\mu} + \rho\mathbf{y}_{t-1} + \mathbf{u}_t, \quad t = -50, \dots, T, \\ \text{DGP2: } \mathbf{y}_t &= \boldsymbol{\mu} + (1 - \rho)\boldsymbol{\beta}t + \rho\mathbf{y}_{t-1} + \mathbf{u}_t, \\ \text{DGP3: } \mathbf{y}_t &= (1 - \rho)\boldsymbol{\mu} + \rho\mathbf{y}_{t-1} + \mathbf{u}_t, \quad \mathbf{u}_t = \mathbf{c} \circ \mathbf{u}_{t-1} + \mathbf{e}_t, \end{aligned}$$

where bold entries indicate vectors of dimension $N \times 1$ and \circ denotes the Hadamard product. The first two DGPs formalize AR(1) models with serially uncorrelated errors, whereas the last one introduces AR(1) disturbances. DGPs 1 and 3 formalize the panel unit root against a panel stationary process with individual effects, while DGP 2 models a panel random walk with drift under H_0 or a panel of trend stationary processes with individual effects under the alternative. In order to account for the deterministic terms (DGPs 1 and 2) and residual serial correlation (DGP3), all tests are computed on the appropriately transformed data as discussed in detail in Chapter 4.3.2 and 4.3.3. Rejection frequencies under H_0 are computed with $\rho = 1$ whereas empirical (size adjusted) power is calculated against the homogeneous alternatives $\rho = 1 - \frac{5}{T\sqrt{N}}$ or $\rho = 1 - \frac{5}{TN^{1/4}}$ for the cases featuring individual intercepts or trends, respectively. Since homogenous PURTs have power against heterogenous alternatives (see Chapter 4.2), it is important

to note that the choice of a homogenous alternative is without loss of generality. Following Pesaran (2007), the deterministic terms are parameterized such that the processes display the same average trend properties under H_0 and the alternative hypothesis. In particular, $\boldsymbol{\mu} \sim iidU(0, 0.02)$, and $\boldsymbol{\beta} \sim iidU(0, 0.02)$. The parametrization of the short run dynamics in DGP 3 is also taken from Pesaran (2007), i.e. $\boldsymbol{c} \sim iidU(0.2, 0.4)$.

Six distinct scenarios for the covariance matrix Ω_t are simulated for each DGP. With regard to contemporaneous correlation, cases of cross sectionally independent, as well as of (weakly) contemporaneously correlated panels are considered. Three different scenarios are simulated with respect to volatility breaks: constant volatility as well as a late positive and an early negative variance shift. Cross sectionally uncorrelated data is generated by setting $\Psi = I_N$ and $\Phi_t = \sigma_{ut}^2 I_N$. The choice of cross sectionally homogenous variances is without loss of generality for the t_{rob} and t_{Wh} statistics but necessary to obtain asymptotic Gaussianity of t_{OLS} in the benchmark case of constant volatility. For the case of a contemporaneously correlated panel, a spatial autoregressive (SAR) error structure is presumed. The latter is specified as

$$\boldsymbol{u}_t = (I_N - \Theta W)^{-1} \boldsymbol{\varepsilon}_t, \quad \text{with } \Theta = 0.8 \quad \text{and} \quad \boldsymbol{\varepsilon}_t \sim iidN(\mathbf{0}, \sigma_{\varepsilon_t}^2 I_N),$$

where the so-called spatial weights matrix W is a row normalized symmetric contiguity matrix of the one-behind-one-ahead type (for more details on spatial panel models see e.g. Elhorst, 2003). In the following, this specification is referred to as an SAR(1) model. The resulting covariance matrix of \boldsymbol{u}_t is given by $\Omega_t = \sigma_{\varepsilon_t}^2 (B' B)^{-1}$ with $B = (I_N - \Theta W)$. As mentioned above, three distinct variance patterns are simulated. Let $\sigma_{u[sT]} = \sigma_{u1} \mathbb{I}(s \leq s_B) + \sigma_{u2} \mathbb{I}(s > s_B)$, where $s_B \in [0, 1]$ indicates the timing of the variance break, $[sT]$ denotes the integer part of sT and \mathbb{I} is the indicator function. In the homoskedastic case, $\sigma_{ut} = \sigma_{u1}$, with $\sigma_{u1} = 1$. The break scenarios are taken from Cavaliere and Taylor (2007b) and are parameterized as $s_B = 0.2$ and $\sigma_{u2} = 1/3$ for the early negative break, while the late positive break is given by $s_B = 0.8$ and $\sigma_{u2} = 3$.

Data is generated for all combinations of $N \in [10, 50]$ and $T \in [10, 50, 100, 250]$. To ensure convergence of the process to its unconditional mean under the alternative hypothesis, 50 presample values are generated and discarded throughout. To compute empirical rejection probabilities under H_0 , each PURT statistic is calculated for the appropriately transformed data and compared with with the 5%

critical value of the Gaussian distribution. Reported estimates for local power are adjusted such that empirical type one errors equal 5%. Throughout, 5000 replications are used.

6.4.2 Results

Table 6.1 documents empirical rejection frequencies obtained for DGP1. The left hand side of Table 1 documents results obtained under cross sectional independence while entries on the right hand side refer to results obtained under a SAR(1) error model. Rejection frequencies under H_0 are reported to the left of size adjusted local power estimates in both cases.

The first block in the upper left panel corresponds to the benchmark case of cross sectional independence and time invariant innovation variances. In this setting, all employed statistics have a Gaussian limiting distribution and, hence, should display empirical rejection frequencies close to 5% as T and N become large. However, the documented results reflect some evidence of small sample size distortions. Empirical rejection frequencies obtained by t_{OLS} range around 7% for panels with $N = 10$, whereas application of t_{rob} leads to undersizing for small values of T . Results obtained for the ‘White-type’ statistic t_{Wh} display comparatively small deviations from the nominal level, especially if $N = 50$. Size adjusted local power estimates indicate that under full homogeneity, all three statistics are asymptotically equally powerful and that the chosen sample sizes are too small for local power estimates to fully converge. The right hand side of the first block presents results for the SAR(1) error model with constant volatility. While the OLS test is severely oversized in this instance, both robust tests remain asymptotically Gaussian. However, finite sample distortions observed for t_{Wh} are slightly larger while the undersizing of t_{rob} is less pronounced than in the case of cross sectional independence. Local power results show that all considered tests are less powerful if the data is cross sectionally correlated. This finding might be explained by noting that cross sectional correlation reduces the amount of independent information contained in the data (Hanck, 2009a).

In line with the theoretical results in Section 6.3.2, results obtained under an early negative variance break and cross sectional independence indicate a tendency of undersizing for t_{OLS} and t_{rob} , where the downward bias of empirical

6.4 Monte Carlo study

Table 6.1: DGP1

		CS independence						SAR(1) model					
		$\rho=1$		$\rho = 1 - \frac{5}{T\sqrt{N}}$				$\rho=1$		$\rho = 1 - \frac{5}{T\sqrt{N}}$			
<i>N</i>	<i>T</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>
<i>Constant variance</i>													
10	10	6.9	2.0	5.8	36.1	35.1	30.0	23.8	4.7	5.8	19.9	18.1	14.9
10	50	6.8	5.0	6.5	39.1	38.6	37.4	25.2	8.2	8.5	20.3	17.0	16.7
10	100	6.9	6.1	6.7	43.8	43.7	44.3	24.4	7.1	7.1	25.0	20.2	20.5
10	250	7.2	6.8	7.2	57.0	56.5	55.8	25.1	8.3	8.5	28.3	21.1	20.9
50	10	5.2	0.0	5.3	47.0	46.9	38.1	20.1	0.5	5.3	21.4	20.5	18.2
50	50	5.8	1.3	5.7	59.0	57.9	57.2	20.8	4.3	7.0	26.8	24.8	23.0
50	100	5.2	2.3	5.3	74.8	74.5	73.3	20.1	4.5	5.9	36.9	33.9	33.5
50	250	5.5	3.9	5.6	84.8	84.4	84.2	21.1	5.7	6.4	44.3	39.4	38.5
<i>Early negative variance shift</i>													
10	10	3.3	0.2	5.8	12.5	13.3	10.2	15.2	1.9	4.6	8.3	7.5	9.0
10	50	4.5	2.7	6.7	9.9	9.7	10.0	19.9	5.8	6.9	8.3	7.4	6.9
10	100	3.8	2.9	6.0	17.3	17.3	14.9	18.7	6.7	7.7	11.0	9.1	8.4
10	250	4.2	3.6	6.1	38.2	38.6	35.6	21.1	7.5	7.9	19.5	15.3	13.5
50	10	1.5	0.0	4.4	8.3	10.1	6.7	13.0	0.0	5.4	7.4	7.8	5.7
50	50	2.8	0.1	5.5	19.6	20.6	17.4	15.4	1.4	5.8	11.9	11.5	11.6
50	100	2.7	0.2	5.5	50.5	49.7	49.4	16.1	2.0	6.6	22.4	20.3	19.2
50	250	3.0	1.5	5.4	92.1	90.8	90.9	15.8	3.1	6.2	49.3	43.1	42.1
<i>Late positive variance shift</i>													
10	10	12.7	4.6	3.5	20.3	20.5	22.3	29.2	8.5	3.9	14.4	13.7	13.5
10	50	13.5	10.8	6.0	24.9	24.1	23.6	31.9	12.0	6.2	19.1	15.3	14.6
10	100	12.7	11.5	6.3	28.6	28.0	26.2	32.7	13.1	6.7	19.8	16.8	14.6
10	250	14.5	13.8	7.1	30.7	30.5	27.0	34.8	14.4	8.2	19.3	14.9	13.5
50	10	11.4	0.1	3.4	22.5	22.7	25.5	26.8	1.9	3.5	12.7	13.3	14.5
50	50	13.0	4.0	5.2	31.6	30.0	31.2	30.2	9.1	5.8	17.8	17.0	16.6
50	100	12.8	8.2	6.3	34.4	32.5	33.0	29.3	11.3	6.3	20.6	19.3	17.8
50	250	12.7	10.1	6.0	40.8	39.6	39.4	30.2	12.5	6.6	21.6	20.4	18.8

Notes: *OLS*, *rob* and *Wh* refer to the PURT statistics defined in (6.4), (6.5),(6.6). Results are based on 5000 replications and the nominal size equals 5%. Local power results are size adjusted. Data is generated according to DGP1 and all statistics are computed on demeaned data as outlined in Chapter 4.3.2.

rejection frequencies positively depends on the size of N . As mentioned before, this is in contrast to results for univariate unit root tests, where positive size distortions are reported (e.g. Kim et al., 2002, Cavaliere and Taylor, 2007b). Rejection frequencies obtained by the ‘White-type’ statistic t_{Wh} display only minor deviations from the nominal significance level. Documented results under spatially correlated errors indicate that size distortions reported for t_{OLS} are less pronounced than under constant volatility since the upward distortion invoked by cross sectional dependence is somewhat dampened by the negative shift in the innovation variance. Empirical rejection frequencies of t_{rob} reflect moderate oversizing for panels with $N = 10$ and $T \geq 50$ and tend to be undersized if $N = 50$. Empirical results for t_{Wh} are only indicative of a moderate finite sample size distortions but are otherwise very similar to those results obtained under constant volatility. With regard to local power, the scenario of an early downward shift in the innovation variance is characterized by a steeper gradient of rejection frequencies with respect to the sample size. While local power estimates are significantly smaller than in the constant variance case for small panel dimensions, up to six percentage points (respectively four percentage points in the SAR(1) case) higher rejection frequencies are documented for the largest simulated panel. The finding of superior power in large samples is supported by the analytically derived location parameter μ_l . Increased asymptotic local power is implied by the absolute value of the location parameter, which becomes larger compared with the benchmark scenario under a downward break in the innovation variance.

If the innovation variance features an upward shift towards the end of the sample, empirical rejection frequencies for t_{OLS} are in the range of 11.4-14.5% for all combinations of N and T and cross sectional independence. Rejection frequencies for t_{rob} depend on the relative magnitude of the time dimension: for T large relative to N , the unit root null hypothesis is rejected significantly too often while for N larger than T , the undersizing observed in the previous experiments persists. Observed upward distortions are in accordance with the theoretical results in Proposition 6.2 and quantitatively in line with results obtained in a similar setting for the univariate DF test (Hamori and Tokihisa, 1997). In contrast, most accurate size control is obtained by t_{Wh} , with empirical errors in rejection frequencies ranging between 0.2 and 2.1 percentage points. If the data is cross sectionally correlated, positive size distortions observed for t_{OLS} and, to

a lesser extent, t_{rob} , are even more pronounced whilst t_{Wh} retains comparatively accurate size control. Results obtained under the alternative hypothesis show that local power estimates are less sensitive to the sample size compared with the case of an early downward shift of innovation variances. However, in line with the asymptotic results in Proposition 6.4, an upward break in the innovation variance induces decreased local power estimates for the largest considered panel dimension.

Table 6.2 reports results for DGP2, with all test statistics computed on detrended data. For the benchmark scenario of constant variances and either cross sectional independence or a SAR(1) error structure, results under H_0 are similar to those obtained for DGP1. As before, a large T relative to N is required in order to obtain rejection probabilities close to 5% for t_{rob} and t_{OLS} yields substantial size distortions under spatial correlation while t_{Wh} provides reliable size control in most instances. Noting that local power is computed in a neighborhood of order $O(T^{-1}N^{-1/4})$, the results imply that local power of all three tests is substantially reduced compared with the intercepts only case of DGP1. For both scenarios of variance shifts, all tests based on detrended data lose size control. If there is a reduction in the innovation variance, the tests are characterized by empirical rejection frequencies which increase with the sample size. In contrast, empirical rejection frequencies of all tests tend to zero in the case of a late positive variance shift. It is noted in Breitung (2000) that the employed detrending scheme is based on the implicit assumption of constant innovation variances. Obviously, the violation of this assumption invokes substantial adverse effects on the performance of the considered PURTs. We do not comment local power results for the latter two scenarios featuring variance shifts, as corresponding size estimates of the tests appear prohibitive for applied research.

Table 6.3 document results for data featuring serially correlated disturbances. These results indicate a general tendency of the tests to overreject H_0 if T is small, with most severe size distortions observed in the case of $N = 50$ and $T = 10$. The latter observation, however, does not apply to t_{rob} , which remains undersized for this panel dimension. Imprecise size estimates for panels with small T are also not surprising from a theoretical point of view. The estimates \hat{c}_i in the prewhitening regression (4.9) are \sqrt{T} consistent and, hence, a relatively large time dimension is required in order to fully remove the effects of serial correlation from

6.4 Monte Carlo study

Table 6.2: Empirical rejection frequencies, DGP2

		CS independence						SAR(1) model					
		$\rho=1$			$\rho = 1 - \frac{5}{TN^{1/4}}$			$\rho=1$			$\rho = 1 - \frac{5}{TN^{1/4}}$		
<i>N</i>	<i>T</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>
<i>Constant variance</i>													
10	10	7.4	1.8	5.4	16.2	16.3	14.5	23.7	5.1	5.4	10.6	10.4	8.9
10	50	6.9	4.9	6.3	17.5	17.3	15.7	22.7	7.0	7.0	11.3	10.1	8.9
10	100	6.1	5.2	5.7	18.7	18.6	19.1	21.0	6.6	6.7	12.0	11.3	11.1
10	250	5.8	5.5	5.7	19.7	19.2	19.3	22.1	6.9	6.9	12.8	11.7	11.8
50	10	6.6	0.0	5.2	22.6	23.3	19.3	21.6	0.7	5.2	12.8	13.2	11.2
50	50	5.6	1.1	5.2	23.6	24.1	22.9	20.9	3.9	6.1	13.8	13.0	12.7
50	100	5.8	2.6	5.6	24.7	24.8	24.3	20.8	4.9	6.5	12.7	12.6	11.6
50	250	5.2	3.5	5.2	28.1	28.3	27.6	21.4	5.4	6.1	15.0	13.7	13.7
<i>Early negative variance shift</i>													
10	10	8.4	0.9	3.9	20.7	19.5	23.7	23.9	4.6	3.3	11.5	11.4	13.5
10	50	14.6	9.7	9.3	13.8	13.8	15.0	32.2	12.6	8.5	8.7	8.0	8.4
10	100	15.4	12.6	11.2	13.2	12.6	13.9	34.0	15.1	10.0	10.2	9.0	8.1
10	250	15.2	14.2	11.4	16.1	15.8	14.8	33.1	15.3	9.9	10.5	9.5	10.0
50	10	11.9	0.0	5.4	32.8	30.9	34.0	26.7	0.1	5.0	15.6	14.8	17.3
50	50	22.1	2.2	16.0	21.9	19.3	25.3	35.8	7.2	10.3	11.5	10.9	12.6
50	100	23.4	7.6	18.3	23.1	21.6	24.6	37.9	11.9	12.6	11.6	10.5	10.6
50	250	23.5	14.3	19.0	30.4	28.5	30.5	38.7	14.8	13.1	14.9	13.8	13.8
<i>Late positive variance shift</i>													
10	10	0.1	0.0	0.0	17.4	15.9	14.5	1.8	0.2	0.1	13.6	13.6	13.2
10	50	0.0	0.0	0.0	18.4	18.7	17.9	2.1	0.4	0.2	12.3	12.0	12.0
10	100	0.0	0.0	0.0	20.1	19.9	20.0	1.9	0.5	0.2	13.2	12.9	13.4
10	250	0.1	0.1	0.1	17.4	17.4	17.2	2.2	0.8	0.5	11.9	11.6	11.4
50	10	0.0	0.0	0.0	25.1	15.2	10.4	0.0	0.0	0.0	14.4	12.7	10.4
50	50	0.0	0.0	0.0	31.2	27.5	18.8	0.0	0.0	0.0	19.2	19.3	18.2
50	100	0.0	0.0	0.0	35.6	31.8	27.6	0.0	0.0	0.0	18.9	19.1	19.2
50	250	0.0	0.0	0.0	35.2	33.4	32.6	0.0	0.0	0.0	18.8	19.3	18.9

Notes: Data is generated according to DGP2 and all tests are computed on detrended data as described in Chapter 4.3.2. For further notes see Table 1.

6.4 Monte Carlo study

Table 6.3: Empirical rejection frequencies, DGP3

		CS independence						SAR(1) model					
		$\rho=1$		$\rho = 1 - \frac{5}{T\sqrt{N}}$				$\rho=1$		$\rho = 1 - \frac{5}{T\sqrt{N}}$			
<i>N</i>	<i>T</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>
<i>Constant variance</i>													
10	10	7.5	1.8	6.0	24.9	24.6	20.1	22.7	3.3	4.4	16.1	14.7	13.0
10	50	7.0	5.0	6.6	36.9	36.9	36.8	24.6	7.5	7.6	19.4	16.6	15.9
10	100	7.1	6.2	6.9	42.0	41.3	41.4	24.4	7.6	7.8	22.2	18.4	18.7
10	250	7.2	6.7	7.0	55.9	55.7	55.4	24.3	8.4	8.6	27.3	21.5	20.9
50	10	14.5	0.0	12.2	27.3	27.6	19.6	28.2	0.2	6.5	16.6	16.1	14.9
50	50	6.8	1.6	6.6	54.9	54.1	53.1	22.6	4.3	6.9	25.1	24.4	23.9
50	100	6.5	2.9	6.5	68.3	68.6	67.5	21.2	5.5	7.0	33.4	30.1	30.0
50	250	5.9	4.4	6.1	83.9	84.3	83.5	21.3	5.8	6.2	43.6	40.3	40.7
<i>Early negative variance shift</i>													
10	10	5.5	0.5	7.9	8.0	8.3	6.2	17.5	2.0	6.4	7.6	7.1	6.2
10	50	4.3	2.7	6.2	10.5	10.6	9.1	18.6	5.5	6.4	8.2	7.2	6.5
10	100	4.4	3.3	7.0	15.6	16.0	14.1	19.2	6.5	7.1	9.5	8.5	9.0
10	250	4.9	4.4	6.9	35.0	34.4	31.7	19.8	7.5	8.1	17.4	12.6	12.5
50	10	9.4	0.0	12.4	4.1	5.6	1.9	22.6	0.0	8.3	5.4	6.0	3.8
50	50	5.0	0.1	9.4	13.9	15.1	12.4	19.0	1.5	7.0	9.6	9.2	9.4
50	100	4.0	0.4	6.9	45.0	45.4	43.1	17.5	2.8	6.7	22.4	19.5	19.0
50	250	3.1	1.3	5.6	91.1	89.6	90.0	16.9	3.4	7.2	45.8	39.7	38.3
<i>Late positive variance shift</i>													
10	10	8.1	1.2	4.9	13.6	14.2	13.4	20.8	2.6	3.9	12.0	12.8	10.5
10	50	14.7	11.7	6.8	22.7	23.2	20.7	33.0	12.3	6.7	15.4	13.4	12.8
10	100	14.1	12.4	6.5	27.3	26.3	24.6	33.2	13.8	7.6	17.2	14.5	13.6
10	250	14.1	13.7	7.4	30.6	30.2	26.0	34.6	14.7	8.1	18.3	14.9	13.8
50	10	10.7	0.0	6.7	13.5	13.6	13.1	23.6	0.2	5.4	11.3	11.1	9.3
50	50	18.8	6.8	8.0	28.6	28.2	28.4	33.9	10.8	7.8	16.0	14.7	13.6
50	100	16.0	9.2	7.1	35.2	33.5	33.3	31.1	11.5	7.0	19.4	17.5	16.6
50	250	14.2	11.5	6.8	39.1	37.4	37.6	31.3	12.9	6.8	21.0	19.3	17.9

Notes: Data is generated according to DGP3 and all tests are computed on prewhitened and centered data, see Chapters 4.3.2 and 4.3.3 for details. For further notes see Table 1.

the data. Conditional on this finding, results obtained under H_0 are qualitatively similar to those obtained for DGP1. In particular, an early negative variance shift diminishes rejection probabilities under H_0 , while a late positive shift leads to increased rejections of H_0 . Moreover, t_{Wh} remains robust against time varying volatility and, as before, application of t_{OLS} leads to markedly oversized rejection rates if the data is cross sectionally correlated. Local power estimates are similar to those obtained for serially uncorrelated error terms (DGP1) with some loss of local power for small values of T .

6.4.3 Summary of simulation results

The main result obtained by the simulation study is that an early negative (late positive) variance shift invokes a downward (upward) distortion of rejection frequencies for PURTs derived under the assumption of invariant second order moments. If the DGP formalizes a random walk without drift under H_0 , rejection rates obtained by the ‘White-type’ statistic t_{Wh} are not affected by variance breaks. Results under the local alternative H_l and the largest considered sample size confirm the theoretical finding that local power is asymptotically higher (lower) under a downward (upward) shift in the innovation variance. However, local power estimates in smaller samples are not necessarily in line with this asymptotic result. For the scenario of a random walk with drift under H_0 , the applied detrending scheme (Breitung, 2000) leads to deceptive inference if there is a break in the innovation variance. Prewhitening the data to remove the effect of serially correlated error terms leaves the main findings unaffected, however, a larger time dimension is required for the empirical type one errors of the tests to come reasonably close to the nominal level.

6.5 Testing the Fisher hypothesis by means of PURTs

6.5.1 Economic background

The Fisher hypothesis (Fisher, 1930) postulates a stable one-to-one relationship between nominal interest rates and the expected rate of inflation. This hypothesis has been investigated in numerous empirical studies (see e.g. Rose, 1988, Crowder, 2003, Cooray, 2003 or Herwartz and Reimers, 2006, 2009). In its simplest form, the Fisher hypothesis states that the nominal interest rate in country i at time t , R_{it} , comprises the ex-ante real interest rate, $E_{t-1}[r_{it}]$, and the ex-ante expected inflation rate, $E_{t-1}[\pi_{it}]$, i.e.

$$R_{it} = E_{t-1}[r_{it}] + E_{t-1}[\pi_{it}] + v_{it},$$

where v_{it} denotes an uninformative forecast error. Under rational expectations, actual and expected values differ only by a white-noise error term, i.e. $\pi_{it} = E_{t-1}[\pi_{it}] + \nu_{it}^{(1)}$ and $r_{it} = E_{t-1}[r_{it}] + \nu_{it}^{(2)}$. Accordingly, the ex-post real interest rate can be expressed as

$$r_{it} = R_{it} - \pi_{it} + \nu_{it}, \quad (6.8)$$

with $\nu_{it} = v_{it} - \nu_{it}^{(1)} - \nu_{it}^{(2)}$. The representation in (6.8) is a starting point for empirical investigations of the Fisher hypothesis by means of unit root tests. If, for instance, inflation and nominal interest rates are found to be I(1) variables, the Fisher hypothesis would imply (1, -1) cointegration establishing a stationary real interest rate. In contrast, a finding of nominal interest rates being I(1) and inflation being I(0) would contradict the Fisher hypothesis.

Prevalence of the Fisher hypothesis is still a question open to empirical research. Using univariate unit root tests on data for 18 economies, Rose (1988) concludes that nominal interest rates follow a unit root process while inflation rates are stationary. On the other hand, Rapach and Weber (2004) report evidence in favor of both variables being integrated of order one, albeit not forming a cointegration relationship. Evidence favorable for a stable long run relationship between inflation and nominal interest rates is reported in Crowder (2003) and Herwartz and Reimers (2006, 2009). However, assessments of the Fisher hypothesis based on first generation PURTs yield conflicting results. For instance,

6.5 Testing the Fisher hypothesis by means of PURTs

Table 6.4: Interest rates, definitions

Country	Label	Interest rate
Belgium	BEL	Treasury paper
Canada	CAN	Treasury Bill rate
France	FRA	Government Bond yield
Germany	GER	Call money rate
Italy	ITA	Government Bond yield medium-term
Japan	JAP	Lending rate
Netherlands	NED	Government Bond yield
United Kingdom	UKD	Treasury Bill rate
United States	USA	Treasury Bill rate

Crowder (2003) finds some evidence of stationary nominal interest rates based on the PURT of Levin et al. (2002) for a panel of 9 industrialized economies. In the latter case, it is argued that these results must be interpreted carefully, as first generation PURTs are generally prone to distorted rejection frequencies through (neglected) cross sectional correlation. Moreover, as highlighted by Kaliva (2008), analyses of the Fisher hypothesis must explicitly account for time-varying volatility as interest and inflation data display marked discrete volatility shifts. In the following assessment of the Fisher hypothesis, the presence of volatility breaks and cross sectional dependence in inflation and interest rate panel data sets is documented. Subsequently, the PURTs discussed above are applied to the data to compare the marginal impacts of accounting for both departures from the assumptions underlying first- and second generation PURTs.

6.5.2 Data and preliminary analyses

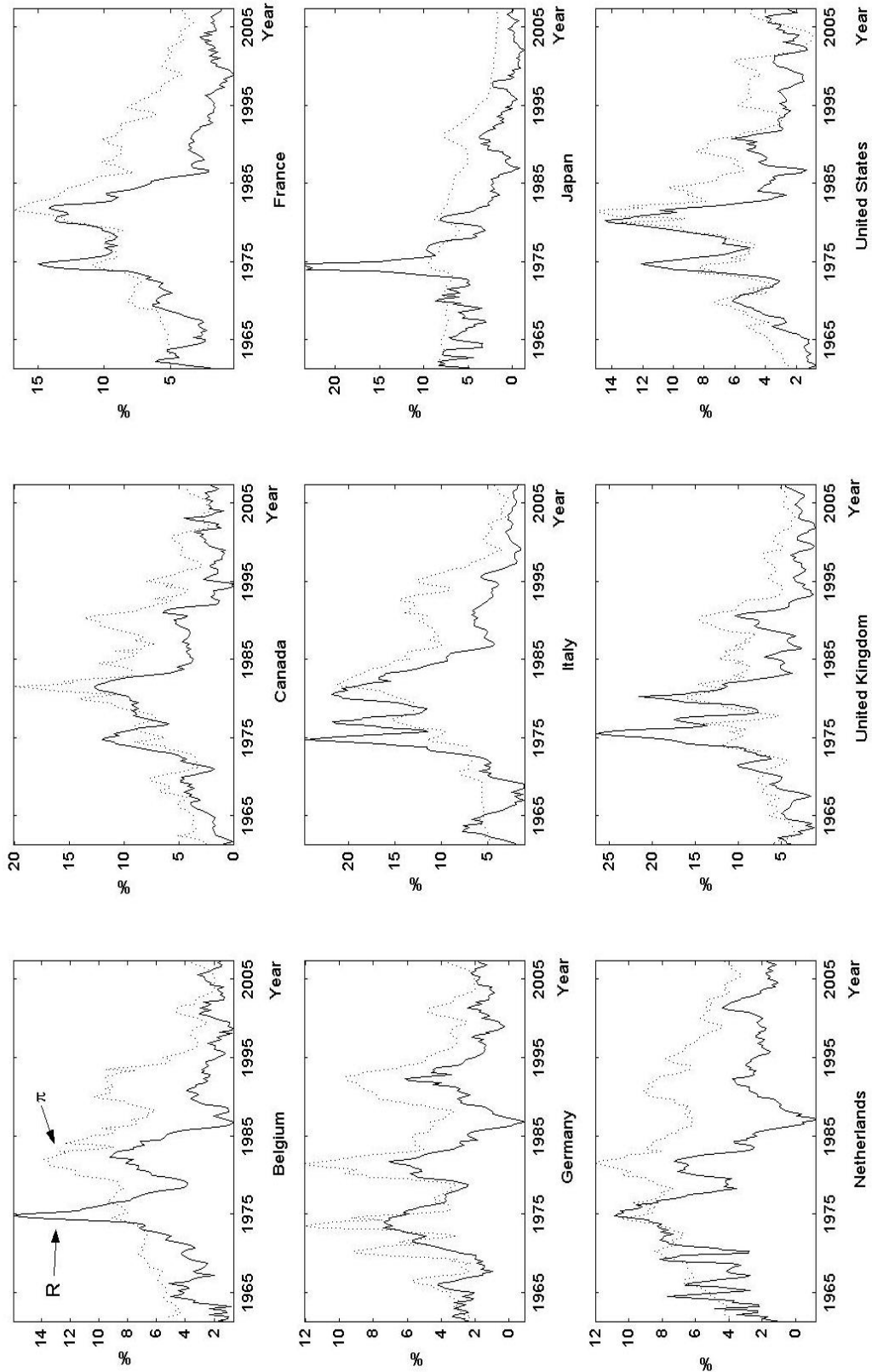
The empirical illustration is based upon the same sample of 9 developed economies considered in Crowder (2003).¹ Data is drawn from the International Financial Statistics of the IMF at the quarterly frequency, ranging from 1961Q2 to 2007Q2.² Inflation rates π_{it} are annual changes of the CPIs. Nominal interest rates, R_{it} , are selected depending on data availability and real interest rates, r_{it} ,

¹These countries are: Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, the United Kingdom and the United States.

²CPI data for the Netherlands is drawn from Dutch national statistics office as IFS data displays discretionary jumps, leading to inflation rates ranging between +30% and -17%.

6.5 Testing the Fisher hypothesis by means of PURTs

Figure 6.2: Nominal Interest Rates and Inflation rates, 1961Q2 - 2007Q2

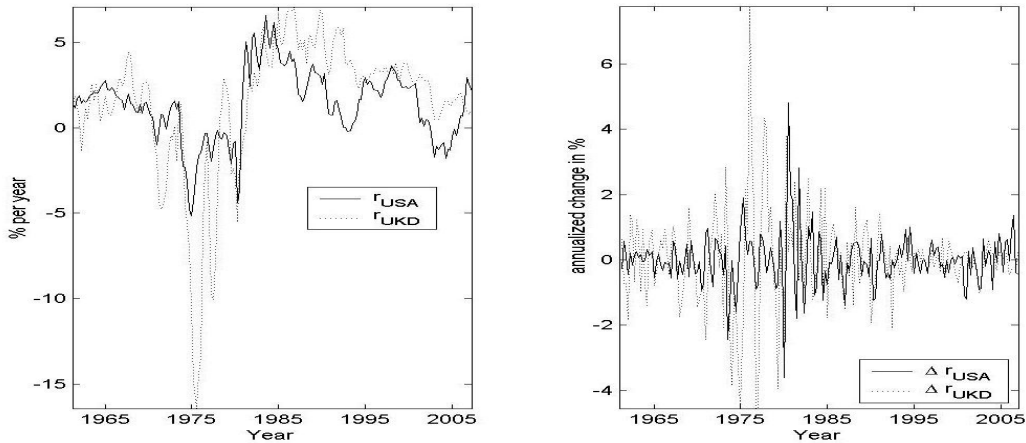


6.5 Testing the Fisher hypothesis by means of PURTs

are obtained as $r_{it} = R_{it} - \pi_{it}$. Table 6.4 contains country specific definitions of interest rate data. The sample data is depicted in Figure 6.2 and eyeball inspection reveals close accordance with the figures provided in Crowder (2003).

Figure 6.3 illustrates the prevalence of cross sectional dependence and time varying volatility. The left hand side graph documents a high degree of comovement of US and UK real interest rates over the sample period. This is not surprising, given that both economies are highly integrated in the world economy and face similar external shocks, as for instance, abrupt oil price swings. The right hand side graph displays the first differences of the two time series, confirming a substantial reduction of volatility around 1985, ending roughly a decade of rather high fluctuations of real interest rates. Crowder (2003).

Figure 6.3: Real interest rates, levels and 1st differences, US vs. UK



The estimated variance profiles $\hat{\vartheta}_i(s)$ of the three variables under investigation are displayed in Figure 6.4 in order to get an impression of the volatility processes governing the sample data (see Cavaliere and Taylor (2007b) for details and alternative estimators of variance profiles). Variance profiles $\vartheta_i(s)$ are calculated as

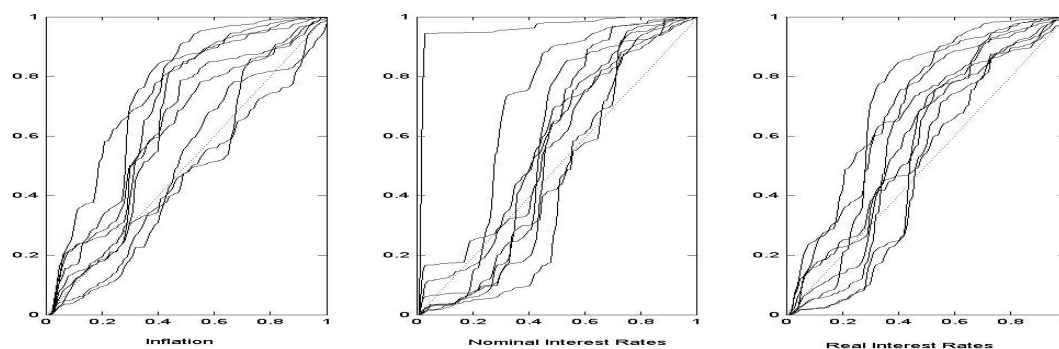
$$\hat{\vartheta}_i(s) = \frac{\sum_{t=1}^{\lfloor sT \rfloor} \hat{e}_{it}^2 + (sT - \lfloor sT \rfloor) \hat{e}_{i[\lfloor sT \rfloor + 1]}^2}{\sum_{t=1}^T \hat{e}_{it}^2}, \quad (6.9)$$

where the \hat{e}_{it} 's are residuals from the first order autoregression of the considered process. While a (perfectly) homoskedastic variance profile would be represented

6.5 Testing the Fisher hypothesis by means of PURTs

by the 45° line, time varying volatilities are characterized by marked deviations from the diagonal.

Figure 6.4: Estimated variance profiles



Inspection of Figure 6.4 reveals that time-varying variances are rather the rule more than an exception for most cross section members. Moreover, it is obvious that estimated variance profiles differ across countries. However, focussing on the overall picture, there is some evidence of an upward followed by a downward shift in the first half of the sample period for all three variables and most of the economies.

In the following, it is analyzed to what extent previous evidence on the Fisher hypothesis obtained via first generation PURTs might have been distorted by cross sectional correlation or (unconditional) volatility shifts.

6.5.3 Panel unit root test results

The first step of the empirical analysis is to prewhiten the raw data. We use the SIC to determine individual specific lag lengths and subsequently apply the prewhitening procedure discussed in Chapter 4.3.3. In order to obtain a balanced panel, the maximum of the individual lag lengths is applied to all cross sectional units, hence prewhitening regressions for most cross sectional units are likely moderately over-fitted. We use 12, 5 and 8 lags of the first differenced series for prewhitening inflation, nominal interest, and real interest rates, respectively. Assuming that inflation as well as interest rates contain a non-zero mean under the stationary alternative, prewhitened data is centered by subtracting the first

6.5 Testing the Fisher hypothesis by means of PURTs

Table 6.5: Empirical results

Variable	T	OLS	rob	Wh
π	172	-3.52 (.000)	-2.45 (.007)	-1.85 (.032)
R	179	-4.22 (.000)	-2.60 (.005)	-1.67 (.048)
r	176	-4.69 (.000)	-3.49 (.000)	-2.83 (.002)

Notes: T denotes the number of included time series observation in the balanced panels. OLS , rob , and Wh refer to the PURT statistics defined in (4.3), (4.14),(5.8). Numbers in parentheses are p -values.

observations. All PURTs are then computed for the resulting balanced panels of prewhitened and centered data. Table 6.5 lists the results of PURT evidence on the Fisher hypothesis. Test statistics for the pooled PURTs are documented in columns 3-5. The numbers in parentheses are p -values obtained from the Gaussian CDF. Results for the three variables are listed by rows.

Using the statistic t_{OLS} to test the order of integration of the inflation rate yields a t -ratio of -3.52 and, hence, a rejection of the unit root null hypothesis at any conventional significance level. This result is in line with Crowder (2003), reporting a t -ratio -5.32 obtained via the Levin et al. (2002) procedure. Given that based on univariate tests, the unit root hypothesis is maintained for all sample economies, Crowder (2003) argues that the rejection of H_0 obtained by the PURT might be due to size bias, invoked by cross sectional dependence. Accordingly, the robust t_{rob} statistic proposed by Breitung and Das (2005) is applied. The resulting t -ratio of -2.45 is substantially smaller in absolute value, however, it still leads to a rejection of the null hypothesis at the 1% significance level. The relative impact of time varying volatility of the sample data on pooled PURTs might be assessed by application of the volatility break robust statistic t_{Wh} . The resulting t -ratio of -1.85 is larger than the t -ratios obtained by t_{OLS} and t_{rob} and the corresponding marginal significance level is 3.2%.

Qualitatively similar results are obtained for the nominal interest rate. By means of the first generation test statistic t_{OLS} , a t -ratio of -4.22 is calculated, which is substantially smaller in absolute value than -7.57 reported in Crowder (2003), but nevertheless leads to a clear rejection of H_0 . Again, application of

the robust tests yields t -ratios which are notably smaller in absolute values. The t -ratio of -2.60 obtained for the cross sectional dependence robust test statistic t_{rob} still implies a rejection of H_0 at the 1% level. However, depending on the chosen nominal significance level, application of t_{Wh} might lead to a different test decision, given the respective p -value of 0.048.

Finally, the unit root hypothesis is tested for the real interest rate. All tests yield results in support of panel stationarity of the real interest rate, and thus, of the Fisher hypothesis. Note however, that at the 5% significance level, even the volatility break robust test does not rule out the possibility of inflation and nominal interest rates being likewise panel stationary variables. Accordingly, one should be careful in interpreting stationarity of real interest rates as a cointegration relationship, linking two nonstationary variables.

6.6 Conclusions

In this chapter, the effects of discrete breaks in the innovation variance on homogenous panel unit root tests are investigated. It is shown that size distortions documented in the literature on univariate unit root tests under time varying variances carry over to the panel case.

The limiting distribution of first and second generation pooled PURTs under a discrete variance shift are derived and it is shown that only the ‘White-type’ PURT statistic proposed in Chapter 5 remains asymptotically Gaussian under the unit root null hypothesis. Under local-to-unity alternatives, it turns out that local power depends on the particular pattern of breaks in the innovation variance. By means of a Monte Carlo study a variety of possible model settings are analyzed, including deterministic trends, autocorrelated disturbances and cross sectional correlation. The simulation study reveals that the ‘White-type’ statistic offers most reliable size control in finite samples and is asymptotically as powerful as the statistic proposed by Breitung and Das (2005). Moreover, it turns out that the employed detrending scheme to account for linear time trends leads to deceptive inference for all analyzed statistics if there is a break in the innovation variance. As an empirical illustration, evidence on the Fisher hypothesis in Crowder (2003) is reconsidered. Based on data for a cross section of 9 developed economies, sampled over the period 1961Q2 - 2007Q2, the order of integration of inflation

rates as well as of nominal and real interest rates is tested. The results illustrate the importance of robust panel unit root tests, accounting for nonstationary innovation variances and cross sectional dependence.

The results in this chapter raise a number of issues for future research. Firstly, noting that the detrending scheme proposed in Breitung (2000) is apparently not applicable under time varying innovation variances, it appears promising to study alternative detrending schemes. Secondly, the assumed constancy of cross sectional correlation might not generally hold in empirical applications. It seems sensible to investigate how time varying patterns of cross sectional correlation affect the performance of PURTs and if the proposed robust statistic is also able to cope with this kind of nuisance appropriately. Finally, the focus of this chapter was on PURTs which are pivotal only under weak cross sectional dependence. Extending the analysis to the case of strong form cross sectional dependence is a topic of immediate interest, which will be covered in the next chapter.

6.7 Appendix

6.7.1 Proof of Proposition 6.1

Basically, all subsequent proofs are extensions of the proofs in Breitung and Das (2005) to the case of discrete variance breaks. To derive the limiting distribution of t_{OLS} define

$$t_{OLS} = \frac{N^{-0.5}T^{-1} \sum_{t=1}^T \mathbf{y}_{t-1} \Delta \mathbf{y}_t}{\sqrt{N^{-1}T^{-2} \sum_{t=1}^T \hat{\sigma}_u^2 \mathbf{y}'_{t-1} \mathbf{y}_{t-1}}} = \frac{a_{NT}}{\sqrt{b_{OLS}}}.$$

Consider the numerator first. Under H_0 , it follows that

$$a_{NT} = N^{-0.5}T^{-1} \sum_{t=1}^T \mathbf{y}_{t-1} \Delta \mathbf{y}_t = N^{-0.5}T^{-1} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{u}_t.$$

Noting that Ω_t can be decomposed as

$$\Omega_t = \begin{cases} \Omega_1 = \Gamma \Lambda_1 \Gamma', & \text{if } 0 < t \leq T_1 \\ \Omega_2 = \Gamma \Lambda_2 \Gamma', & \text{if } T_1 < t \leq T. \end{cases},$$

where $\Lambda_\bullet = \text{diag}(\lambda_1, \dots, \lambda_N)'$, $\bullet = 1, 2$, is a diagonal matrix of eigenvalues and Γ is the corresponding matrix of normalized eigenvectors, which remains unaffected by the shift in idiosyncratic variance components due to the assumed time invariant pattern of cross sectional correlation. Now that $\mathbf{e}_t = \Lambda_\bullet^{-1/2} \Gamma' \mathbf{u}_t$ is an $N \times 1$ vector of cross sectionally independent error terms with unit variance and $\mathbf{z}_t = \Lambda_\bullet^{-1/2} \Gamma' \mathbf{y}_t$, is an $N \times 1$ vector of mutually uncorrelated random walks, the numerator can be expressed as

$$\begin{aligned} a_{NT} &= N^{-0.5}T^{-1} \left[\sum_{t=1}^{T_1} \left(\sum_{s=1}^{t-1} \mathbf{e}_s \right)' \Gamma \Lambda_1 \Gamma' \mathbf{e}_t \right. \\ &\quad \left. + \sum_{t=T_1+1}^T \left(\sum_{s=1}^{T_1} \mathbf{e}_s \right)' \Gamma \Lambda_1^{1/2} \Lambda_2^{1/2} \Gamma' \mathbf{e}_t \sum_{T_1+1}^t \left(\sum_{s=T_1+1}^{t-1} \mathbf{e}_s \right)' \Gamma \Lambda_2 \Gamma' \mathbf{e}_t \right], \end{aligned} \quad (6.10)$$

The terms in (6.10) are constructed such that summation always only comprises error terms with homogenous variances as, for instance, \mathbf{e}_s is a multivariate Gaussian random vector. It holds accordingly that $T_1^{-1/2} \sum_{s=1}^{T_1} \mathbf{e}_s = \mathbf{z}_{T_1} \xrightarrow{d} \mathbf{W}(1)$,

where $\mathbf{W}(r)$ is a multivariate standard Brownian motion. Defining $\bar{\mathbf{z}}_{t-1} = \mathbf{z}_{t-1} - \mathbf{z}_{T_1}$, one obtains

$$a_{NT} = N^{-0.5}T^{-1} \left[\sum_{i=1}^N \lambda_{1i} \sum_{t=1}^{T_1} z_{i,t-1} e_{it} + \sum_{i=1}^N \lambda_{1i}^{1/2} \lambda_{2i}^{1/2} z_{iT_1} \sum_{t=T_1+1}^T e_{it} + \sum_{i=1}^N \lambda_{2i} \sum_{t=T_1+1}^T \bar{z}_{i,t-1} e_{it} \right].$$

To economize on space, the shorthand notations $\int W_i$ and $\int W_i dW_i$ instead of $\int W_i(r)dr$ and $\int W_i(r)dW_i(r)$ are used in the following. As $T, T_1 \rightarrow \infty$, common invariance principles for partial sum processes imply that

$$a_{NT} \xrightarrow{d} N^{-0.5} \left[\delta \sum_{i=1}^N \lambda_{1i} \int_0^1 W_i dW_i + \sqrt{\delta(1-\delta)} \sum_{i=1}^N \sqrt{\lambda_{1i}\lambda_{2i}} W_{i,T_1}(1) W_{i,T_2}(1) + (1-\delta) \sum_{i=1}^N \lambda_{2i} \int_0^1 W_i dW_i \right], \quad (6.11)$$

where δ is defined as in \mathcal{A}_2 . The subscripts in $W_{i,T_1}(1)$ and $W_{i,T_2}(1)$ in the medium term of the right hand side of (6.11) are chosen in order to highlight that both terms are the values of two uncorrelated Brownian motions at $r = 1$ with $T_2 = T - T_1$. Since $W_{i,T_1}(1)$ and $W_{i,T_2}(1)$ are independent Gaussian random variables and $E \left[\int_0^1 W_i dW_i \right] = 0$ while $Var \left[\int_0^1 W_i dW_i \right] = 0.5$, one obtains from the central limit theorem for mean zero *iid* random variables that the numerator of the three test statistics t_{OLS} , t_{rob} , and t_{Wh} is given by

$$a_{NT} \xrightarrow{d} N(0, \bar{\sigma}^2), \quad \bar{\sigma}^2 = 0.5\delta^2 \bar{\lambda}_1^2 + \delta(1-\delta) \bar{\lambda}_1 \bar{\lambda}_2 + 0.5(1-\delta)^2 \bar{\lambda}_2^2, \quad (6.12)$$

where $\bar{\lambda}_\bullet^2 = N^{-1} \sum_{i=1}^N \lambda_\bullet^2$, with $\bullet = 1, 2$, and $\bar{\lambda}_1 \bar{\lambda}_2 = N^{-1} \sum_{i=1}^N \lambda_1 \lambda_2$ as $N \rightarrow \infty$.

Now consider the denominator of t_{OLS} . We have

$$b_{OLS} = N^{-1}T^{-2}\hat{\sigma}^2 \sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{y}_{t-1} = N^{-1}T^{-2}\hat{\sigma}^2 \left[\sum_{t=1}^{T_1} \mathbf{z}'_{t-1} \Lambda_1 \mathbf{z}_{t-1} + T_1 T_2 \frac{\mathbf{z}'_{T_1}}{\sqrt{T_1}} \Lambda_1 \frac{\mathbf{z}_{T_1}}{\sqrt{T_1}} + \sum_{t=T_1+1}^T \bar{\mathbf{z}}'_{t-1} \Lambda_2 \bar{\mathbf{z}}_{t-1} \right].$$

As $T \rightarrow \infty$,

$$b_{OLS} \xrightarrow{d} N^{-1} \left(N^{-1} \delta \sum_{i=1}^N \lambda_{1i} + N^{-1} (1 - \delta) \sum_{i=1}^N \lambda_{2i} \right) \\ \times \left[\delta^2 \sum_{i=1}^N \lambda_{1i} \int_0^1 W_i^2 + \delta(1 - \delta) \sum_{i=1}^N \lambda_{1i} W_i(1)^2 + (1 - \delta)^2 \sum_{i=1}^N \lambda_{2i} \int_0^1 W_i^2 \right].$$

Letting $N \rightarrow \infty$, convergence in probability follows

$$b_{OLS} \xrightarrow{p} (\delta \bar{\lambda}_1 + (1 - \delta) \bar{\lambda}_2) [0.5 \delta^2 \bar{\lambda}_1 + \delta(1 - \delta) \bar{\lambda}_1 + 0.5(1 - \delta)^2 \bar{\lambda}_2], \quad (6.13)$$

since $E[\int_0^1 W_i^2] = 0.5$ and $E[W_i(1)^2] = 1$. It is immediate from (6.13) that $b_{OLS} \neq \bar{\sigma}^2$, implying that t_{OLS} does not converge to a Gaussian limiting distribution if there is a break in the innovation variance, even under cross sectional independence and cross sectionally homogeneous variances.

□

6.7.2 Proof of Proposition 6.2

Since the numerator is the same for t_{OLS} , t_{rob} , and t_{Wh} , it suffices to consider the denominator to derive the asymptotic distribution of t_{rob} . Specifically,

$$b_{rob} = \sum_{t=1}^T \mathbf{y}'_{t-1} \hat{\Omega} \mathbf{y}_{t-1}, \text{ with } \hat{\Omega} = T^{-1} \sum_{t=1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}'_t = T^{-1} \sum_{t=1}^T \mathbf{u}_t \mathbf{u}'_t + o_p(1).$$

Making use of the same decomposition as in (6.10) and dropping lower order terms yields

$$b_{rob} = N^{-1} T^{-2} \left[\sum_{t=1}^{T_1} \mathbf{z}'_{t-1} \Lambda_1^{1/2} \left\{ \frac{T_1}{T} \Lambda_1 \left(T_1^{-1} \sum_{t=1}^{T_1} \mathbf{e}_t \mathbf{e}'_t \right) + \frac{T_2}{T} \Lambda_2 \left(T_2^{-1} \sum_{t=T_1+1}^T \mathbf{e}_t \mathbf{e}'_t \right) \right\} \Lambda_1^{1/2} \mathbf{z}_{t-1} \right. \\ + T_1 T_2 \frac{\mathbf{z}'_{T_1}}{\sqrt{T_1}} \Lambda_1^{1/2} \left\{ \frac{T_1}{T} \Lambda_1 \left(T_1^{-1} \sum_{t=1}^{T_1} \mathbf{e}_t \mathbf{e}'_t \right) + \frac{T_2}{T} \Lambda_2 \left(T_2^{-1} \sum_{t=T_1+1}^T \mathbf{e}_t \mathbf{e}'_t \right) \right\} \Lambda_1^{1/2} \frac{\mathbf{z}_{T_1}}{\sqrt{T_1}} \\ \left. + \sum_{t=T_1+1}^T \mathbf{z}'_{t-1} \Lambda_2^{1/2} \left\{ \frac{T_1}{T} \Lambda_1 \left(T_1^{-1} \sum_{t=1}^{T_1} \mathbf{e}_t \mathbf{e}'_t \right) + \frac{T_2}{T} \Lambda_2 \left(T_2^{-1} \sum_{t=T_1+1}^T \mathbf{e}_t \mathbf{e}'_t \right) \right\} \Lambda_2^{1/2} \mathbf{z}_{t-1} \right].$$

As $T \rightarrow \infty$ and by noting that $T^{-1} \sum_{t=1}^T \mathbf{e}_t \mathbf{e}'_t$, $T_1^{-1} \sum_{t=1}^{T_1} \mathbf{e}_t \mathbf{e}'_t$ and $T_2^{-1} \sum_{t=T_1+1}^T \mathbf{e}_t \mathbf{e}'_t \xrightarrow{p} E[\mathbf{e}_t \mathbf{e}'_t] = I_N$, one obtains

$$\begin{aligned} b_{rob} &\xrightarrow{d} N^{-1} \left[\delta^3 \sum_{i=1}^N \lambda_{1i}^2 \int_0^1 W_i^2 + \delta^2(1-\delta) \sum_{i=1}^N \lambda_{1i} \lambda_{2i} \int_0^1 W_i^2 \right. \\ &\quad + \delta^2(1-\delta) \sum_{i=1}^N \lambda_{1i}^2 W_i(1)^2 + \delta(1-\delta)^2 \sum_{i=1}^N \lambda_{1i} \lambda_{2i} W_i(1)^2 \\ &\quad \left. + (1-\delta)^2 \delta \sum_{i=1}^N \lambda_{1i} \lambda_{2i} \int_0^1 W_i^2 + (1-\delta)^3 \sum_{i=1}^N \lambda_{2i}^2 W_i^2 \right]. \end{aligned}$$

For $N \rightarrow \infty$ this yields

$$\begin{aligned} b_{rob} &\xrightarrow{p} 0.5 \left\{ \delta^3 \overline{\lambda_1^2} + (\delta^2(1-\delta) + (1-\delta)^2 \delta) \overline{\lambda_1 \lambda_2} + (1-\delta)^3 \overline{\lambda_2^2} \right\} \\ &\quad + \left\{ \delta^2(1-\delta) \overline{\lambda_1^2} + \delta(1-\delta)^2 \overline{\lambda_1 \lambda_2} \right\} \\ &= (\delta^2 - 0.5\delta^3) \overline{\lambda_1^2} + (1.5\delta + 0.5\delta^2 - 2\delta^3) \overline{\lambda_1 \lambda_2} + 0.5(1-\delta)^3 \overline{\lambda_2^2}, \end{aligned}$$

establishing that

$$\nu_{rob} = \frac{0.5\delta^2 \overline{\lambda_1^2} + \delta(1-\delta) \overline{\lambda_1 \lambda_2} + 0.5(1-\delta)^2 \overline{\lambda_2^2}}{(\delta^2 - 0.5\delta^3) \overline{\lambda_1^2} + (1.5\delta + 0.5\delta^2 - 2\delta^3) \overline{\lambda_1 \lambda_2} + 0.5(1-\delta)^3 \overline{\lambda_2^2}} \neq 1.$$

The result in Breitung and Das (2005) with $\nu_{rob} = 1$ holds as a special case if $\delta = 1$ or $\delta = 0$.

□

6.7.3 Proof of Proposition 6.3

Finally, it is shown that $b_{Wh} \xrightarrow{p} \bar{\sigma}^2$ for $T \rightarrow \infty$ followed by $N \rightarrow \infty$. With $\check{\mathbf{u}}_t = \mathbf{u}_t + o_p(1)$ and dropping the lower order term in the expression it follows that

$$\begin{aligned} b_{Wh} &= N^{-1} T^{-2} \sum_{t=1}^T \mathbf{y}'_{t-1} \mathbf{u}_t \mathbf{u}'_t \mathbf{y}_{t-1} = N^{-1} T^{-2} \sum_{t=1}^T \mathbf{z}'_{t-1} \Lambda \mathbf{e}_t \mathbf{e}'_t \Lambda \mathbf{z}_{t-1} \\ &= N^{-1} T^{-2} \left[\sum_{t=1}^{T_1} \mathbf{z}'_{t-1} \Lambda_1 \mathbf{e}_t \mathbf{e}'_t \Lambda_1 \mathbf{z}_{t-1} + T_1 T_2 \frac{\mathbf{z}'_{T_1}}{\sqrt{T_1}} \Lambda_1^{1/2} \Lambda_2^{1/2} \mathbf{e}_t \mathbf{e}'_t \Lambda_1^{1/2} \Lambda_2^{1/2} \frac{\mathbf{z}_{T_1}}{\sqrt{T_1}} \right. \\ &\quad \left. + \sum_{t=T_1+1}^T \mathbf{z}'_{t-1} \Lambda_2 \mathbf{e}_t \mathbf{e}'_t \Lambda_2 \mathbf{z}_{t-1} \right]. \end{aligned}$$

By Assumption $\mathcal{A}_1(iii)$, restricting $E[e_{it}^4] < \infty$, define $\xi_{it} = e_{it}^2$, which is an *iid* random variable with $E[\xi_{it}] = 1$ and $Var[\xi_{it}] = \sigma_{\xi_{it}}^2 < \infty$. Hence, as $T \rightarrow \infty$,

$$b_{Wh} \xrightarrow{d} N^{-1} \left[\delta^2 \sum_{i=1}^N \lambda_{1i}^2 \int_0^1 W_i^2 \xi_{it} + \delta(1-\delta) \sum_{i=1}^N \lambda_{1i} \lambda_{2i} W_i(1)^2 \xi_{it} + (1-\delta)^2 \sum_{i=1}^N \lambda_{2i}^2 \int_0^1 W_i^2 \xi_{it} \right].$$

Because of the independence of the ξ_{it} and the partial sum processes, as $N \rightarrow \infty$ this expression converges in probability

$$b_{Wh} \xrightarrow{p} 0.5\delta^2 \overline{\lambda_1^2} + \delta(1-\delta) \overline{\lambda_1 \lambda_2} + 0.5(1-\delta)^2 \overline{\lambda_2^2} = \overline{\sigma^2},$$

verifying that $t_{Wh} \xrightarrow{d} N(0, 1)$.

□

6.7.4 Proof of Proposition 6.4

The derivation of the limiting distribution of t_{Wh} under the sequence of local alternatives $H_l : \rho = 1 - \frac{c}{T\sqrt{N}}$ is based on the respective proof for the statistic t_{rob} in Breitung and Das (2005). First note that in local-to-unity neighborhoods as defined above, $\mathbf{z}_{[rT]} \xrightarrow{d} \mathbf{W}_i(r)$ for all $0 \leq c < \infty$. It follows that the numerator of t_{Wh} is given by

$$a_{Wh} = T^{-1} N^{-1/2} \sum_{t=1}^T \mathbf{y}_{t-1} \Delta \mathbf{y}_t = T^{-1} N^{-1/2} \sum_{t=1}^T \mathbf{e}'_t \Lambda \Delta \mathbf{z}_t - c T^{-2} N^{-1} \sum_{t=1}^T \mathbf{z}'_{t-1} \Lambda \mathbf{z}_{t-1}.$$

From the proof of Proposition 6.7.1 it follows directly that the first term on the right hand side equals the numerator under the null hypothesis while the second term converges in probability

$$c T^{-2} N^{-1} \sum_{t=1}^T \mathbf{z}'_{t-1} \Lambda \mathbf{z}_{t-1} \xrightarrow{p} c \left[0.5\delta^2 \overline{\lambda_1} + \delta(1-\delta) \overline{\lambda_1} + 0.5(1-\delta)^2 \overline{\lambda_2} \right],$$

as $T \rightarrow \infty$, followed by $N \rightarrow \infty$. From the proof of Proposition 6.3 it follows that $t_{Wh} \xrightarrow{d} N(-c\mu_l, 1)$, with

$$\mu_l = \frac{0.5\delta^2\bar{\lambda}_1 + \delta(1-\delta)\bar{\lambda}_1 + 0.5(1-\delta)^2\bar{\lambda}_2}{\sqrt{0.5\delta^2\bar{\lambda}_1^2 + \delta(1-\delta)\bar{\lambda}_1\bar{\lambda}_2 + 0.5(1-\delta)^2\bar{\lambda}_2^2}}.$$

Again, the result in Breitung and Das (2005) with $\mu_{l,rob} = \sqrt{0.5\bar{\lambda}_1}/\sqrt{\bar{\lambda}_1^2}$ obtains as a special case with $\delta = 0$ or $\delta = 1$.

□

Chapter 7

Testing convergence of unit labor costs in the Euro area: nonstationary volatility and wild bootstrap panel unit root tests

7.1 Introduction

Long run convergence of inflation differentials among European Economic and Monetary Union (EMU) member economies is considered a prerequisite for the union's long term viability. Persistent inflation differentials not only cast doubt on the adequacy of a common "one size fits all" monetary policy, but also imply growing dispersion of intra EMU price competitiveness. As an eventual consequence, high inflation economies might consider pulling out of the EMU in order to restore their external competitiveness by means of a nominal devaluation. It is a common approach in the literature to test convergence of inflation rates by means of univariate and panel unit root tests (e.g. Caporale and Pittis, 1993, Beck and Weber, 2005, Acosta, 2007, Buseti et al., 2007 and Lopez and Papell, 2008

among others). Most authors find evidence in favor of inflation convergence prior to the introduction of the Euro, while the period following the Euro introduction is characterized by more persistent or even diverging inflation differentials. A notable exception is the work of Lopez and Papell (2008) who are able to identify a faster rate of convergence after the advent of the Euro. However, according to Balassa (1964) and Samuelson (1964), persistent CPI inflation differentials may constitute an equilibrium phenomenon as lower income economies are often characterized by higher average inflation during the process of real convergence. Due to asymmetric productivity developments across sectors, positive inflation differentials do not necessarily impair international price competitiveness, even within a monetary union. To account for this effect, convergence of unit labor costs (ULC) growth differentials has been considered as an productivity adjusted alternative to testing for convergence of CPI inflation (e.g. Fischer, 2007, Dullien and Fritsche, 2008 or Fritsche and Kuzin, 2007).

The contribution of this chapter is twofold. Noting that time series data on ULC inflation differentials displays marked shifts in volatility, standard univariate and panel unit root tests (PURT) lose asymptotic pivotalness. There is a growing literature on unit root tests under nonstationary volatility (e.g. Hamori and Tokihisa, 1997, Kim et al., 2002, Cavaliere, 2004 or Cavaliere and Taylor, 2007a,b, 2008). In contrast, the effects of nonstationary volatility on homogenous PURTs have not yet been discussed in the literature. As shown in Chapter 6, breaks in the unconditional second moment of the data generating process lead to distortions of empirical rejection frequencies of first (Levin et al., 2002) and second generation (Breitung and Das, 2005) homogenous PURTs. On the other hand, the ‘White’ corrected statistic proposed in Chapter 5.3.1 retains its Gaussian limiting distribution. However, this robust test statistic allows only for weak form cross sectional dependence as defined by Breitung and Pesaran (2008) and requires a large cross sectional dimension for asymptotic Gaussianity. Both assumptions appear overly restrictive in the present application. First, the

7.2 PURTs under nonstationary volatility

co-evolution of EMU economies' ULC inflation rates suggests the presence of strong form cross sectional dependence. Moreover, the considered data set is characterized by a small cross sectional dimension. Wild bootstrap variants of homogenous PURTs as introduced in Chapter 5.5 adequately cope with a small cross sectional dimension as well as with strong form cross sectional dependence. In this chapter, it is shown that this bootstrap approach remains valid even under general forms of nonstationary volatility. Secondly, it is demonstrated that conclusions on ULC convergence are indeed sensitive with respect to the application of bootstrap critical values. Based on the bootstrap approach, there is some significant evidence in favor of diverging ULC inflation dynamics within the EMU since the introduction of the Euro.

7.2 PURTs under nonstationary volatility

7.2.1 The autoregressive, heteroskedastic panel model

The following discussion is based on the autoregressive panel model

$$\mathbf{y}_t = \rho \mathbf{y}_{t-1} + \mathbf{u}_t, \quad t = 1, \dots, T. \quad (7.1)$$

As before, $\mathbf{y}_t = (y_{1t}, \dots, y_{Nt})'$, $\mathbf{y}_{t-1} = (y_{1,t-1}, \dots, y_{N,t-1})'$ and $\mathbf{u}_t = (u_{1t}, \dots, u_{Nt})'$ are $N \times 1$ vectors and the index $i = 1, \dots, N$ indicates cross sectional units. The AR coefficient ρ satisfies either $\rho = 1$ under the unit root null hypothesis or $|\rho| < 1$ under the stationary alternative hypothesis. The following set of assumptions is made with respect to the vector of disturbances \mathbf{u}_t :

Assumption 7.1 (\mathcal{A}_1)

- (i) The disturbance vector $\mathbf{u}_t = \Omega_t^{1/2} \mathbf{e}_t$, $\mathbf{e}_t \sim iid(\mathbf{0}, I_N)$.
- (ii) The volatility matrix $\Omega_t^{1/2}$ is non-stochastic and satisfies $\Omega_{\lfloor rT \rfloor}^{1/2} := \Omega^{1/2}(r)$ with $\lfloor rT \rfloor$ denoting the integer part of rT for all $r \in [0, 1]$. Moreover, $\Omega(r)$ is

assumed to be positive definite for all $r \in [0, 1]$.

(iii) Finally, it is assumed that $E[u_{it}u_{jt}u_{kt}u_{lt}] < \infty$ for all i, j, k, l .

The assumptions $\mathcal{A}_1(i)$ - $\mathcal{A}_1(ii)$ are the same as in Cavaliere et al. (2010). In particular, $\mathcal{A}_1(ii)$ allows for general forms of time varying nonstationary volatility, such as discrete jumps, trending volatility or smooth transition variance shifts. As there are no further restrictions on the covariance matrix apart from positive definiteness, strong form cross sectional dependence as defined by Breitung and Pesaran (2008) is explicitly allowed for. Thus, by allowing more general forms of cross sectional dependence as well as a broad class of (nonstochastic) volatility processes, the assumptions above are less restrictive than \mathcal{A}_1 and \mathcal{A}_2 considered in the previous chapter.

7.3 Wild bootstrap PURTs

It has been shown in the previous Chapter that both the first generation statistic t_{OLS} and the second generation statistic t_{rob} lose asymptotic pivotalness under discrete variance breaks. Even though the test statistic t_{Wh} is adequately immunized against nonstationary volatility and (weak) cross sectional dependence, there are situations in which inference based on bootstrap critical values might be preferable. For instance, asymptotic Gaussianity of t_{Wh} requires T and N approaching infinity. However, in many macroeconomic applications availability of monthly data is often restricted to the rather small sample of highly developed industrial economies or even only a subgroup thereof. As shown in Chapter 5.5, wild bootstrap PURTs are valid under a finite cross sectional dimension and thus are a viable alternative in such applications. Moreover, under strong form cross sectional dependence, t_{Wh} is not pivotal. Converging to the same (non-standard) asymptotic distribution, wild bootstrap PURTs allow for correct inference in these particular cases. Finally, it is well known (e.g. Horowitz,

2001) that resampling of pivotal test statistics can yield asymptotic refinements, i.e. faster convergence of the test's actual size to the nominal significance level than obtained by asymptotic approximations. It is demonstrated in Cavaliere and Taylor (2008) that wild bootstrap M -unit root tests are robust with respect to a wide class of nonstationary volatility. Hence, it appears straightforward to extend this approach to the panel case.

7.3.1 The bootstrap algorithm

Recall the wild bootstrap algorithm introduced in Chapter 5.5. Wild resampling of homogenous PURTs proceeds as follows:

1. Run the pooled DF regression in (5.2) and obtain OLS residuals $\hat{\mathbf{u}}_t = \Delta \mathbf{y}_t - \hat{\phi} \mathbf{y}_{t-1}$. Calculate the corresponding PURT statistic, denoted ψ .
2. Replicate sufficiently often the following steps:
 - (i) draw bootstrap residuals \mathbf{u}_t^* from $\hat{\mathbf{u}}_t$ as

$$\mathbf{u}_t^* = (u_{1t}^*, u_{2t}^*, \dots, u_{Nt}^*)' = \eta_t (\hat{u}_{1t}, \hat{u}_{2t}, \dots, \hat{u}_{Nt})', \quad \eta_t \sim iid(0, 1),$$

where η_t , $t = 1, \dots, T$, is independent from the panel data;

- (ii) construct the bootstrap sample \mathbf{y}_t^* according to the DGP presumed under H_0 as

$$\mathbf{y}_t^* = \mathbf{y}_{t-1}^* + \mathbf{u}_t^*, \quad \mathbf{y}_0^* = \mathbf{y}_0;$$

- (iii) calculate the bootstrap version ψ^* of ψ .

3. Decision: Reject H_0 with significance α if $\psi < c_\alpha^*$, the α -quantile of ψ^* .

Several choices of η_t are available from the literature (Liu, 1988; Mammen, 1993). While Cavaliere and Taylor (2008) use the Gaussian distribution to obtain wild bootstrap variants of univariate M -unit root tests, Davidson and Flachaire

(2001) highlight the particular merits of the Rademacher distribution. Unlike in Chapter 5, the Gaussian distribution is used to generate the bootstrap sample in the following. An unreported set of simulation shows that the choice of the particular distribution has only a minor impact on the finite sample results. With regard to the construction of $\hat{\mathbf{u}}_t$, Cavaliere and Taylor (2008) point out that in the univariate case $\hat{\mathbf{u}}_t$ obtained by imposing the null hypothesis, i.e. $\hat{\mathbf{u}}_t = \check{\mathbf{u}}_t = \Delta \mathbf{y}_t$ dominates other alternatives in finite samples.

Owing to η_t being serially uncorrelated, the bootstrap sample is also serially uncorrelated even if (possibly serially correlated) pseudo residuals $\check{\mathbf{u}}_t = \Delta \mathbf{y}_t$ are used. Accordingly, there is no need of accounting for serial correlation in the bootstrap sample. If deterministic terms are presumed in the DGP and the data is detrended according to Breitung (2000), the bootstrap sample has to be detrended in the same way.

7.3.2 Asymptotic properties of wild bootstrap PURTs

In this section, asymptotic validity and consistency of the wild bootstrap PURTs under nonstationary volatility is established. To facilitate the analytical derivations, the bootstrap variant of t_{OLS} is considered.

7.3.2.1 Asymptotic validity

Asymptotic validity of the wild bootstrap PURTs under cross sectional dependence as derived in Chapter 5.5.2 relies on the validity of a bootstrap invariance principle in analogy to the multivariate invariance principle for partial sum processes of serially independent increments. However, this invariance principle no longer applies under nonstationary volatility. A multivariate invariance principle and its bootstrap counterpart under nonstationary volatility are established in

Cavaliere et al. (2010). In particular,

$$S_T(r) := T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \mathbf{u}_t \xrightarrow{d} M(r), \quad (7.2)$$

where $M(r) := \int_0^r \Omega(s)^{1/2} dW(s)$ is an N -variate continuous martingale with volatility matrix $\Omega(s)^{1/2}$ and $W(s)$ is an N -variate standard Brownian motion.

The corresponding bootstrap invariance principle is given by

$$S_T^*(r) := T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \mathbf{u}_t^* \xrightarrow{d_p} M(r), \quad (7.3)$$

where $\xrightarrow{d_p}$ denotes weak convergence in probability. In analogy to the proof in Chapter 5.5.2 and using shorthand notation, the limiting distribution of the test statistic can be written as

$$t_{OLS} \xrightarrow{d} \left(\bar{\omega}^2 \sum_{i=1}^N \sum_{j=1}^N \int_0^1 M_i M_j \right)^{-1/2} \left(\sum_{i=1}^N \int_0^1 M_i dM_i \right) \equiv \Xi, \quad (7.4)$$

with $\bar{\omega}^2 = \frac{1}{N} \sum_{i=1}^N \bar{\omega}_i^2$ and $\bar{\omega}_i^2$ is the asymptotic average innovation variance of cross section unit i . Obtaining from Cavaliere and Taylor (2008) weak convergence in probability of the bootstrap residual variance estimator for the i th cross sectional unit $(\hat{\sigma}_u^*)_i^2 \xrightarrow{d_p} \bar{\omega}_i^2$ with $(\hat{\sigma}_u^*)_i^2 = T^{-1} \sum_{t=1}^T (\hat{u}_{it}^*)^2$, and owing to (7.3) it holds accordingly that, conditional on the data,

$$t_{OLS}^* \xrightarrow{d_p} \left(\bar{\omega}^2 \sum_{i=1}^N \sum_{j=1}^N \int_0^1 M_i M_j \right)^{-1/2} \left(\sum_{i=1}^N \int_0^1 M_i dM_i \right). \quad (7.5)$$

Given the asymptotic representations in (7.4) and (7.5) weak convergence in probability of the bootstrap statistic t_{OLS}^* to the limiting distribution Ξ of the original statistic t_{OLS} can be established.

Proposition 7.1 *Let $\mathbf{y}_t = (y_{1t}, \dots, y_{Nt})'$ be a panel AR(1) process, generated according to (7.1) and the Assumptions in \mathcal{A}_1 hold. Then, under $H_0 : \rho_i = 1, \forall i$, with N fixed and as $T \rightarrow \infty$*

$$\sup_{-\infty < g < \infty} |P(t_{OLS}^* \leq g \mid \mathcal{Y}_T) - P(\Xi \leq g)| \xrightarrow{p} 0,$$

where t_{OLS}^* is the wild bootstrap variant of t_{OLS} in (4.3), $\mathcal{Y}_T = (\mathbf{y}_0, \mathbf{u}_1, \dots, \mathbf{u}_T)$ and Ξ is the limit distribution in (7.4).

7.3.2.2 Consistency

To derive the consistency of the bootstrap tests, let $p_{T,OLS}^*$ denote the bootstrap p -value corresponding to t_{OLS}^* . The following Proposition states the consistency of the wild bootstrap statistic under fixed alternatives with $|\rho| < 1$.

Proposition 7.2 *Under H_1 with $|\rho| < 1$ and $T \rightarrow \infty$, $t_{OLS}^* \xrightarrow{d} \Xi$, with Ξ as defined in 7.4. Moreover, t_{OLS} diverges to minus infinity as $T \rightarrow \infty$. Hence, $p_{T,OLS}^* \xrightarrow{p} 0$.*

The intuition behind Proposition 7.2 consists of two parts. The first is to show that the bootstrap statistic retains the same limiting distribution under the alternative hypothesis while the second part is to derive divergence of the original statistic. Fortunately, such arguments are readily available in the literature. Since the bootstrap sample is constructed under the null hypothesis, it suffices to prove consistency of the variance estimator under the alternative hypothesis. This is verified for the univariate case in Cavaliere and Taylor (2008) (proof of their Lemma 3), namely $(1/T) \sum_{t=1}^{\lfloor rt \rfloor} \widehat{u}_t^2 \xrightarrow{p} \int_0^r \omega(s)^2 ds$ uniformly for all $s \in [0, 1]$. As the OLS variance estimator in (4.3) is given as the arithmetic average of the N idiosyncratic components, consistency carries over to the panel case. Divergence of the original statistics under the alternative hypothesis is, for instance, derived in Breitung and Westerlund (2009). In fact, as demonstrated by Breitung and Westerlund (2009), t_{OLS} is also consistent in the case of a heterogeneous alternative hypothesis with $|\rho_i| < 1$ holding only for some non-zero fraction of the cross sectional units.

7.4 Monte Carlo study

7.4.1 The simulation design

The finite sample performance of the wild bootstrap PURTs are investigated by considering model (7.1). This data generating process (DGP) formalizes a highly stylized panel AR(1) model without deterministic terms and serially uncorrelated errors. Concentrating on such a simplistic model allows to highlight the isolated and combined effects of strong cross sectional dependence and non-stationary volatility on the considered PURTs. As will become evident in the next section, empirical data on ULC growth convergence among EMU economies features both, a high degree of cross sectional correlation as well as significant variance breaks. Moreover, in testing ULC growth convergence, the alternative hypothesis is usually formulated as a mean-zero stationary process. Rejection frequencies under H_0 are computed with $\rho = 1$ whereas empirical (size adjusted) power is calculated against the homogenous fixed alternative $\rho = 0.98$. In order to separate the issues of cross sectional dependence and nonstationary volatility, the covariance matrix of the disturbances \mathbf{u}_t can be written as $\Omega_t = \Phi_t^{1/2} \Psi \Phi_t^{1/2}$, where Ψ is a time invariant correlation matrix and $\Phi_t^{1/2}$ is a time varying volatility matrix. Cross sectionally uncorrelated data with $\Psi = I_N$ and $\Phi_t^{1/2} = \sigma_{ut} I_N$ is simulated in order to investigate the isolated effects of nonstationary volatility. The choice of cross sectionally homogenous variances is without loss of generality for the t_{rob} and t_{Wh} statistics but necessary to obtain asymptotic Gaussianity of t_{OLS} in the benchmark scenario of constant volatility. Contemporaneous correlation is modeled by an equicorrelated error structure, where the correlation matrix satisfies $\Psi_{ii} = 1$ and $\Psi_{ij} = 0.8$ for $i \neq j$. In the framework of Breitung and Pesaran (2008), this error structure falls into the category of strong cross sectional dependence as the largest eigenvalue of Ψ increases without bounds as $N \rightarrow \infty$. Three different scenarios are simulated with respect to volatility breaks. The con-

stant volatility case serves as a benchmark against which the break scenarios are compared. In modeling variance breaks, only those cases are considered which yield most significant size distortions in the univariate case, namely a late positive as well as an early negative discrete variance shift. These breaks are modeled as follows: Let $\sigma_{u[rT]} = \sigma_{u1}\mathbb{I}(r \leq r_B) + \sigma_{u2}\mathbb{I}(r > r_B)$, where $r_B \in [0, 1]$ indicates the timing of the variance break and \mathbb{I} is the indicator function. In the homoskedastic case, $\sigma_{ut} = \sigma_{u1}$, with $\sigma_{u1} = 1$. The break scenarios are parameterized as in Cavaliere and Taylor (2007b), thus $r_B = 0.2$ and $\sigma_{u2} = 1/3$ for the early negative break, while the late positive break is given by $r_B = 0.8$ and $\sigma_{u2} = 3$. Data is generated for all combinations of $N \in [5, 10, 50]$ and $T \in [50, 100, 250]$. These combinations cover most relevant situations for macroeconometric applications, where cross section units are usually countries or states and data is sampled at most at the monthly frequency. To allow for non-zero, random initial conditions, 50 presample values are generated and discarded throughout. To compute empirical rejection frequencies under H_0 , the resulting PURT statistics are compared with the 5% critical value of the Gaussian distribution. Reported power estimates are adjusted such that empirical type one errors equal 5%. Throughout, 5000 replications and 499 bootstrap iterations are used.

7.4.2 Results

Tables 7.1 and 7.2 document the results of the simulation study. Empirical rejection frequencies under H_0 are documented in Table 7.1 while size adjusted power is reported in Table 7.2. To enhance readability of the tables, rejection frequencies under H_0 which are not covered by the 95% confidence interval around the nominal level (4.4-5.6%) are highlighted by bold font.

Under H_0 , the results in Table 7.1 for the restrictive benchmark scenario with constant variance and cross sectional independence are indicative of some finite sample distortions for all considered test statistics. Largest deviations from the

7.4 Monte Carlo study

Table 7.1: Finite sample properties: Rejection frequencies under H_0

N	T	CS independence						Constant correlation					
		OLS	rob	Wh	OLS^*	rob^*	Wh^*	OLS	rob	Wh	OLS^*	rob^*	Wh^*
<i>Constant variance</i>													
5	50	6.9	6.8	6.6	5.1	5.1	5.4	23.9	9.1	8.2	5.0	5.5	6.1
5	100	7.2	6.9	7.1	5.1	5.1	5.2	25.5	8.8	8.6	5.1	5.6	6.1
5	250	8.1	7.7	7.8	5.1	5.1	5.2	26.6	9.0	8.7	4.6	5.1	5.3
10	50	6.6	6.1	6.5	5.1	5.5	5.6	35.9	9.0	8.2	5.3	5.8	6.2
10	100	7.2	6.7	7.1	5.4	5.7	5.6	38.3	8.8	8.3	5.0	5.1	5.4
10	250	6.7	6.4	6.7	5.0	5.1	4.9	40.9	8.8	8.4	4.9	5.0	5.2
50	50	5.4	3.0	5.1	4.7	6.8	5.0	54.4	9.5	8.5	4.6	5.9	6.6
50	100	5.8	3.8	5.6	5.2	7.1	5.2	54.3	9.3	9.0	5.6	5.5	5.8
50	250	6.1	4.7	5.9	5.3	6.8	5.3	56.9	9.3	9.2	5.0	5.3	5.4
<i>Early negative variance shift</i>													
5	50	8.5	8.9	6.6	5.8	5.4	6.2	21.8	12.9	6.6	5.4	5.2	5.8
5	100	7.9	8.1	6.3	4.8	4.7	5.1	22.7	13.1	7.0	5.3	5.3	5.4
5	250	8.6	8.6	7.5	5.5	5.6	5.7	26.2	14.8	9.3	6.3	6.2	6.5
10	50	7.6	7.9	6.6	5.9	5.2	6.3	29.5	13.3	7.2	5.2	5.6	6.3
10	100	6.8	6.9	6.4	5.6	5.3	5.9	31.8	12.8	7.0	4.8	4.6	5.6
10	250	6.2	6.0	6.3	5.1	5.0	5.1	36.4	14.3	8.7	4.8	4.9	5.4
50	50	5.7	4.3	5.5	5.4	6.1	5.7	48.6	13.9	7.2	5.1	6.0	6.3
50	100	5.4	3.2	5.3	4.9	6.8	5.3	49.8	14.5	7.4	4.6	5.3	5.6
50	250	4.4	3.2	5.3	5.0	6.8	5.0	54.3	14.1	8.3	4.7	4.9	5.1
<i>Late positive variance shift</i>													
5	50	9.4	8.7	7.1	5.8	6.0	6.0	29.8	9.8	8.2	5.4	6.1	6.4
5	100	10.2	9.6	6.9	5.0	5.3	5.5	31.4	10.6	7.8	5.1	5.1	5.6
5	250	12.7	12.2	7.9	5.0	5.3	5.5	35.3	12.3	8.3	5.0	5.1	5.2
10	50	8.7	7.9	6.5	5.9	6.5	6.0	43.0	9.5	8.0	5.8	5.8	6.1
10	100	10.0	9.3	6.4	4.8	5.5	5.2	45.6	9.8	7.8	4.5	4.8	5.4
10	250	11.3	10.9	6.4	4.8	5.3	5.1	46.7	13.3	8.7	5.7	5.6	5.8
50	50	8.2	4.7	5.6	5.3	8.7	5.5	57.5	9.8	7.9	5.1	6.2	6.1
50	100	9.3	6.0	5.0	4.9	7.7	4.7	59.2	11.5	8.4	5.6	5.9	5.8
50	250	11.4	9.3	5.8	5.2	7.0	5.1	60.2	13.7	9.0	5.1	5.4	5.5

Notes: OLS , rob and Wh refer to the PURT statistics defined in (4.3), (4.14),(5.8). Results are based on 5000 replications and 499 bootstrap repetitions. The nominal size equals 5%. Data is generated according to (7.1).

nominal significance level with errors in rejection probability (ERP) of around 3 percentage points can be observed for a small cross sectional dimension ($N = 5$) and a large time dimension ($T = 250$). While ERPs for t_{OLS} and t_{Wh} decrease for increasing cross sectional dimensions, the time dimension is required to be substantially larger than the cross sectional dimension to obtain rejection frequencies close to the nominal level for t_{rob} . Otherwise, with N larger or of similar magnitude as T , rejection probabilities for t_{rob} are too low. In comparison, the bootstrap statistics offer rather accurate rejection frequencies. Only t_{rob}^* displays four significant deviations from the nominal level whereas t_{OLS}^* and t_{Wh}^* yield rejection frequencies statistically indistinguishable from the nominal level for all sample sizes. This result highlights asymptotic refinements provided by using bootstrap critical values instead of asymptotic approximations. Under the presumed model of cross sectional dependence, the three PURT statistics fail asymptotic pivotalness and, hence, rejection frequencies are substantially upward distorted. However, while rejection frequencies obtained for t_{OLS} diverge with an increasing cross sectional dimension, ERPs of around four percentage points are observed for t_{rob} and t_{Wh} , invariant with respect to the size of the cross sectional dimension. As postulated by the theoretical results, the bootstrap statistics are robust against strong form cross sectional dependence with reported rejection frequencies very close to the nominal level throughout.

Size adjusted power estimates displayed in Table 7.2 illustrate the consistency of all three tests as the power estimates increase along both sample dimensions in the benchmark case. The equality of the power of the original and bootstrap tests derived by Cavaliere and Taylor (2008) is confirmed by the results. If the data is cross sectionally dependent, empirical power drops significantly, a result which is well known in the PURT literature (see e.g. Hanck, 2009a). Moreover, size adjusted power estimates display a clear ordering with t_{OLS} being substantially more powerful than t_{rob} and t_{Wh} . Depending on the sample size, power differentials of up to 30 percentage points can be observed. However, it

7.4 Monte Carlo study

Table 7.2: Finite sample properties: Rejection frequencies under H_1

N	T	CS independence						Constant correlation					
		<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS*</i>	<i>rob*</i>	<i>Wh*</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>OLS*</i>	<i>rob*</i>	<i>Wh*</i>
<i>Constant variance</i>													
5	50	43.1	40.2	41.2	39.8	40.5	39.8	22.6	15.0	14.9	24.3	16.2	14.4
5	100	76.8	75.6	76.0	75.4	75.8	74.9	42.6	24.1	23.8	42.4	24.6	22.4
5	250	100.	99.9	99.9	99.9	99.9	99.9	84.6	57.9	57.8	84.8	58.6	57.8
10	50	73.3	70.6	71.3	72.4	71.5	69.5	28.1	16.6	16.7	31.2	16.6	16.4
10	100	98.5	98.3	98.2	98.3	98.5	98.0	50.5	27.7	27.0	51.4	27.7	26.8
10	250	100.	100.	100.	100.	100.	100.	91.0	62.2	61.9	90.7	62.5	61.2
50	50	100.	100.	100.	100.	100.	100.	34.4	18.7	17.6	42.5	19.4	17.7
50	100	100.	100.	100.	100.	100.	100.	58.0	28.2	26.6	60.1	27.8	26.8
50	250	100.	100.	100.	100.	100.	100.	93.9	62.5	62.2	94.5	62.1	62.1
<i>Early negative variance shift</i>													
5	50	68.2	61.7	65.9	64.9	61.7	64.1	28.0	21.2	23.8	37.4	25.2	23.8
5	100	93.5	90.1	87.7	94.7	92.7	85.7	43.4	27.7	30.1	52.7	29.6	29.8
5	250	100.	100.	98.6	100.	100.	98.1	80.8	49.5	39.9	86.6	53.9	38.3
10	50	96.5	90.9	95.1	96.0	93.7	94.4	32.5	23.4	26.5	47.0	26.8	25.7
10	100	100.	99.9	100.	100.	100.	99.8	58.2	34.3	32.9	67.7	37.0	31.8
10	250	100.	100.	100.	100.	100.	100.	91.5	60.0	48.8	95.1	64.0	47.7
50	50	100.	100.	100.	100.	100.	100.	35.9	26.7	28.7	57.0	26.8	27.9
50	100	100.	100.	100.	100.	100.	100.	63.1	33.7	34.1	79.1	37.4	34.7
50	250	100.	100.	100.	100.	100.	100.	95.7	58.6	50.0	98.2	64.6	50.0
<i>Late positive variance shift</i>													
5	50	25.6	24.2	19.9	21.6	21.3	20.1	21.7	13.3	10.4	17.7	12.2	10.7
5	100	46.9	45.3	33.9	42.3	42.3	34.3	33.3	22.0	17.4	29.9	20.6	17.6
5	250	83.6	83.2	66.7	81.8	81.5	66.8	61.9	47.2	32.7	58.7	43.2	32.8
10	50	39.9	36.3	33.6	35.7	35.1	32.2	25.2	14.3	13.1	21.0	13.3	13.0
10	100	67.8	65.3	56.3	62.6	62.2	55.2	41.3	25.0	19.0	36.5	22.0	19.1
10	250	97.4	97.2	93.1	96.9	96.5	92.3	65.9	44.8	33.3	64.4	42.5	32.6
50	50	91.7	85.3	87.0	88.9	86.6	86.7	32.5	15.3	12.4	27.0	14.0	12.1
50	100	99.7	99.2	99.3	99.5	99.1	99.2	45.1	23.5	18.7	40.4	19.9	18.4
50	250	100.	100.	100.	100.	100.	100.	71.1	48.1	34.4	70.0	46.9	34.1

Notes: See Table 7.1.

has to be noted that empirical rejection frequencies for t_{OLS} obtained under H_0 are severely distorted, hence limiting the applicability of this statistic in applied research. Noting that the bootstrap variant t_{OLS}^* has similar power features as the original statistic t_{OLS} underpins the relative usefulness of the wild bootstrap approach: Besides yielding most accurate size control, the bootstrap variant t_{OLS} offers a substantial power improvement over the alternative statistics.

If the DGP features a negative volatility shift, rejection frequencies under H_0 display only minor differences compared with the benchmark case under cross sectional independence. For a large cross sectional dimension, rejection frequencies for t_{OLS} and t_{rob} tend to decrease below the nominal level. Under cross sectional dependence, the ERPs for t_{rob} are somewhat larger than in the baseline scenario with constant volatility. These results are somewhat at odds with the theoretical propositions and simulation results presented in Chapter 6. However, the difference between these two sets of simulations is the treatment of the deterministic terms, respectively the initial conditions. In the previous chapter, all tests are performed on demeaned data such that the first (transformed) observation equals zero. This is in accordance with the theoretical derivations, in which zero initial conditions are presumed to obtain convergence of the partial sum processes to Brownian motions. However, in this case, the tests are computed under the assumed absence of deterministic terms but non-zero initial values are permitted. Hence, the theoretical results which imply underrejections of H_0 under a downward volatility shift not necessarily apply in finite samples as the starting values might be too far away from the zero line. Unreported results for DGPs imposing $\mathbf{y}_0 = (0, \dots, 0)'$ are in line with the results in Chapter 6.4.2 and indicate that undersizing is much more pronounced in that case. Irrespective of the treatment of the initial values, the bootstrap statistics yield reliable size control - even under cross sectional dependence. More pronounced differences compared with the case of constant volatility can be identified with respect to size adjusted power results. In line with theoretical propositions in Chapter 6, size adjusted power increases

7.5 Persistent inflation differentials and stability of the EMU

compared with the benchmark scenario. The power gain is most substantial for small sample sizes and more pronounced under cross sectional independence.

In line with the theoretical propositions, a positive volatility shift induces upward distortions of empirical rejection frequencies for t_{OLS} and t_{rob} while t_{Wh} retains asymptotic size control and only displays minor finite sample distortions. If the DGP incorporates cross sectional dependence and a positive volatility break, the size distorting effect of (neglected) strong form cross sectional dependence are further exacerbated for t_{OLS} and t_{rob} . In contrast, rejection frequencies based on bootstrap critical values remain very close to the nominal 5% significance level. Among the bootstrap tests, t_{OLS}^* offers most accurate size control, whereas some notable ERPs of up to 3.7 percentage points can be identified for t_{rob}^* . A positive variance break invokes a substantial decline in the empirical power of the tests.

7.5 Persistent inflation differentials and stability of the EMU

7.5.1 Background

Persistent inflation differentials within single economies are usually a matter of benign neglect. It is assumed that factor mobility, in particular labor mobility, eventually equilibrates unsustainable regional inflation disparities. However, in the EMU which, from an economic point of view, can be understood as a single economy, diverging inflation rates are a matter of concern. Whilst intra EMU capital mobility is arguably very high, the European labor market is still highly segmented along national borders. Therefore, the existence of (mean reverting) intra union inflation differentials is explicitly welcomed by the ECB as an important macroeconomic adjustment mechanism (ECB, 2005). Moreover, according to the so-called Balassa-Samuelson effect, long lasting differentials of CPI inflation within a monetary union might also constitute an equilibrium phenomenon.

7.5 Persistent inflation differentials and stability of the EMU

It has been argued by Balassa (1964) and Samuelson (1964) that catching up effects in per capita income lead to persistently higher inflation in fast growing, less developed economies compared with economies closer to their steady state. On the other hand, however, permanent inflation differentials imply accumulating real exchange rate changes, leading to shifts in the relative competitiveness of member economies. Continued divergence of competitive positions inducing (or exacerbating) trade imbalances might lead to political tension within the union. Even though failure of the monetary union is an extreme scenario, stable real exchange rates are a prerequisite for the success of a “one size fits all” common monetary policy. Accordingly, converging inflation levels, i.e. mean-zero stationary inflation differentials are a necessary and sufficient condition of real exchange rates stability within a monetary union.

Inflation convergence is often tested in terms of absolute β -convergence in the framework of Barro and Sala-i-Martin (1991) and Bernard and Durlauf (1996). Absolute β -convergence of inflation rates implies that the sequence of inflation differentials $\{y_t^{ij}\}_{t=1}^T$ between economies i and j is a mean-zero $I(0)$ variable. If y_t^{ij} were a stationary variable with a non-zero mean, relative β -convergence would hold. However, relative β -convergence of inflation rates implies constant inflation differentials and, hence, a secular trend in the competitive position between economies i and j . In order to account for the Balassa-Samuelson effect, ULC growth in the manufacturing sector constitutes a productivity adjusted inflation measure which can be used to test convergence (or rather stability) of competitive positions within the EMU. Dullien and Fritsche (2008) Fritsche and Kuzin (2007) and Fischer (2007) analyze the dynamics of EMU ULC growth differentials using annual, respectively quarterly data. Using first and second generation PURTs computed over the entire sample period, the convergence hypothesis is accepted in these three studies. However, while Fischer (2007) and Fritsche and Kuzin (2007) apply second generation PURTs accounting for cross sectional dependence, the issue of nonstationary volatility has been neglected so far. Furthermore, tests

7.5 Persistent inflation differentials and stability of the EMU

have only been applied to the entire sample period. As the introduction of the Euro likely constitutes a structural break in the underlying economic processes, it appears reasonable to test the convergence hypothesis not only for the whole sample but also for subsamples, spanning over the periods prior to and after the introduction of the Euro. Hereby, it is attempted to shed some more light on the question if inflationary discipline has been loosened after the adoption of the single currency. Busetti et al. (2007) provide some evidence for divergence of CPI inflation differentials after the introduction of the Euro. They argue that the Exchange Rate Mechanism and the Maastricht criteria for the adoption of the Euro strengthened the convergence process prior to the introduction of the Euro, while divergent wage developments have led to increasing disparities thereafter. Here, it is attempted to investigate relative competitive positions among Euro member states by analyzing ULC growth differentials based on the manufacturing sector.

7.5.2 Data set and methodology

Monthly relative normalized ULC indexes (base period 2000 = 100) are compiled via Datastream from the IMF's international financial statistics and are constructed as a ratio of an index of actual hourly compensation per worker in the manufacturing sector by the normalized index of output per man-hour. By using data from the manufacturing sector, empirical results on ULC inflation differentials can be directly interpreted as the relative evolution of external competitiveness within the EMU. The data set is spanning from January 1978 to March 2009, thus we have a total of 375 time series observations at hand, which cover the whole period from stage 1 of the European Monetary System until the current regime of the EMU with the Euro as a common currency. Monthly data is not available for all economies belonging to the "Euro12" group.¹ The consid-

¹These are: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain

7.5 Persistent inflation differentials and stability of the EMU

ered sample consists of eight EMU members, namely Austria, Belgium, Finland, France, Germany, Ireland, Italy and Spain. Among the economies not included in the data set, Portugal and Greece typically display above average inflation rates compared with other Euro economies (Busetti et al., 2007). This might arguably bias the results in favor of the convergence hypothesis. ULC growth differentials are computed as follows: Let c_{it} denote the ULC index for economy i at time t . The annual ULC growth rate in percentage points is then given as

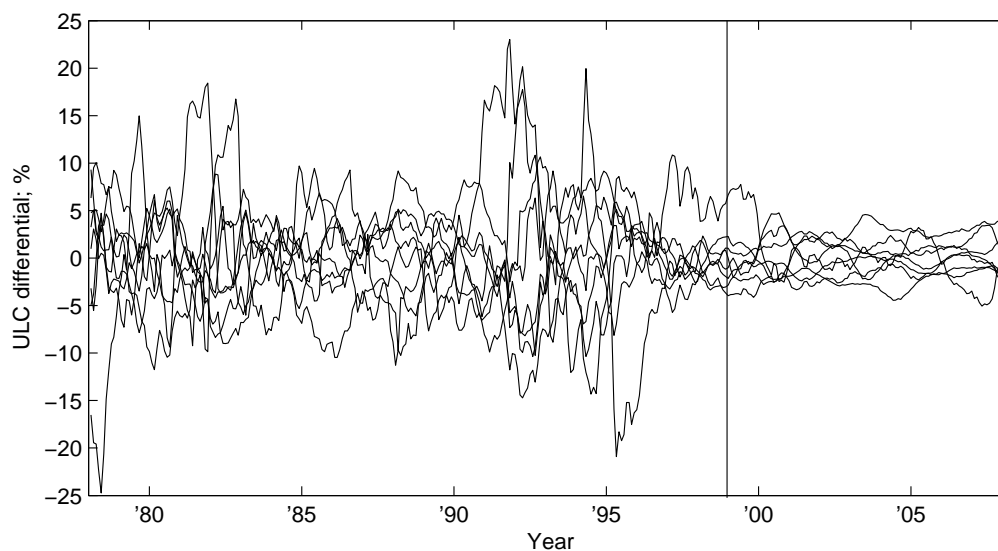
$$\pi_{it} = 100(\ln(c_{it}) - \ln(c_{i,t-12})).$$

Accordingly, 12 initial observations are lost. ULC growth differentials between two economies i and j are then calculated as $y_t^{ij} = \pi_{it} - \pi_{jt}$. Two issues are analyzed. First, joint convergence among the sample economies towards a common ULC growth level is tested. Therefore, PURTs are applied to a panel of ULC growth differentials computed against the cross sectional average as $\tilde{y}_t^i = \pi_{it} - \bar{\pi}_t$, where $\bar{\pi}_t = N^{-1} \sum_{i=1}^N \pi_{it}$. This approach is also pursued by Beck and Weber (2005) testing convergence of CPI inflation among regions of six EMU members. Since rejection of the null hypothesis does not necessarily imply convergence of all time series included in the panel, it is subsequently attempted to identify economies which are characterized by diverging ULC growth dynamics vis-à-vis the remaining sample economies. Both panel and univariate unit root tests are used to assess this second issue.

Figure 7.1 depicts the sample of ULC growth differentials with respect to the cross sectional average. The vertical solid line corresponds to the introduction of the Euro as a single (accounting) currency. It is immediate from the figure that the introduction of the Euro coincides with a massive drop in volatility of ULC growth differentials. While the period prior to the advent of the Euro is characterized by large deviations of up to 20 percentage points from the sample mean (by construction the zero line), deviations from the cross sectional mean are much smaller since the introduction of the Euro. To further analyze the dynamics

7.5 Persistent inflation differentials and stability of the EMU

Figure 7.1: ULC growth differentials vis-a-vis the EU8 average



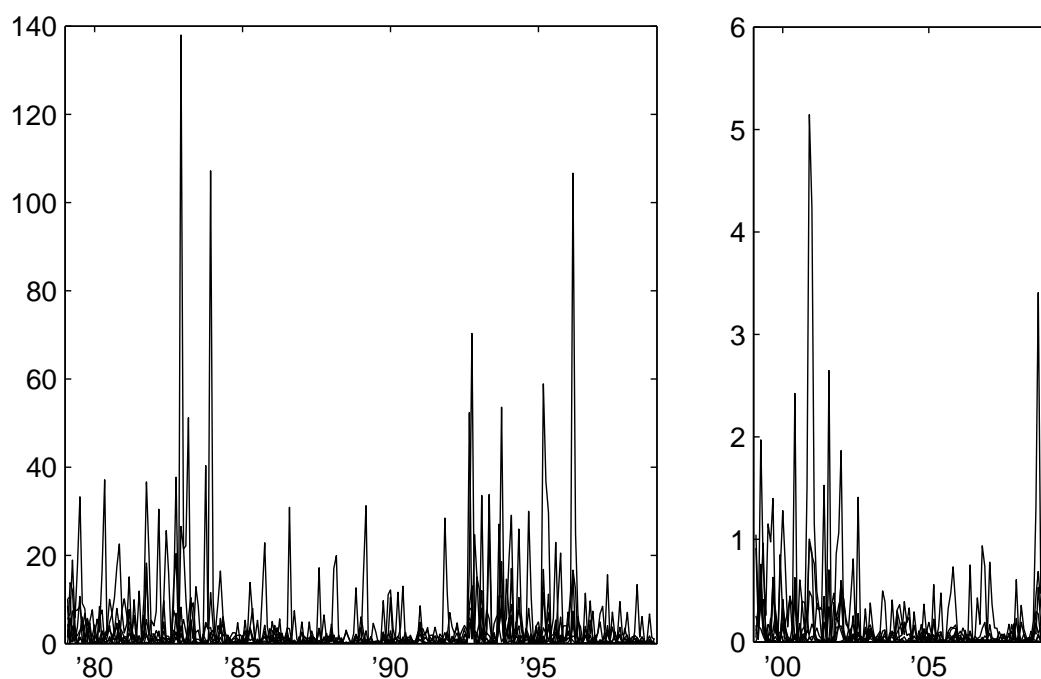
of the data's time varying variance, Figure 7.2 displays the squared differences of the differentials depicted in Figure 7.1, separated into a pre and post Euro introduction subsample.

The different scales on the ordinates underline the substantial decline in ULC growth differentials' volatility in the recent subperiod. This finding implies that the ECB has been relatively successful in anchoring inflationary expectations at a low level. However, there is considerable variation of volatility within these two subsamples. In the first subsample, there is a marked reduction of volatility in the middle of the eighties in accordance with the Great Moderation hypothesis. This period of rather low ULC inflation volatility ends in the aftermath of the European exchange rate crisis in 1992/93, with higher volatility levels prevailing roughly until 1996. For the second subsample an early drop in volatility can be observed. These findings of unstable volatility call for the application of nonstationary volatility robust (panel) unit root tests.

Since identification of economies with (potentially) diverging ULC growth is of major economic interest, country specific analyses complementing the tests on group-wise convergence are additionally required. Figure 7.3 shows panels of

7.5 Persistent inflation differentials and stability of the EMU

Figure 7.2: Volatility of the data

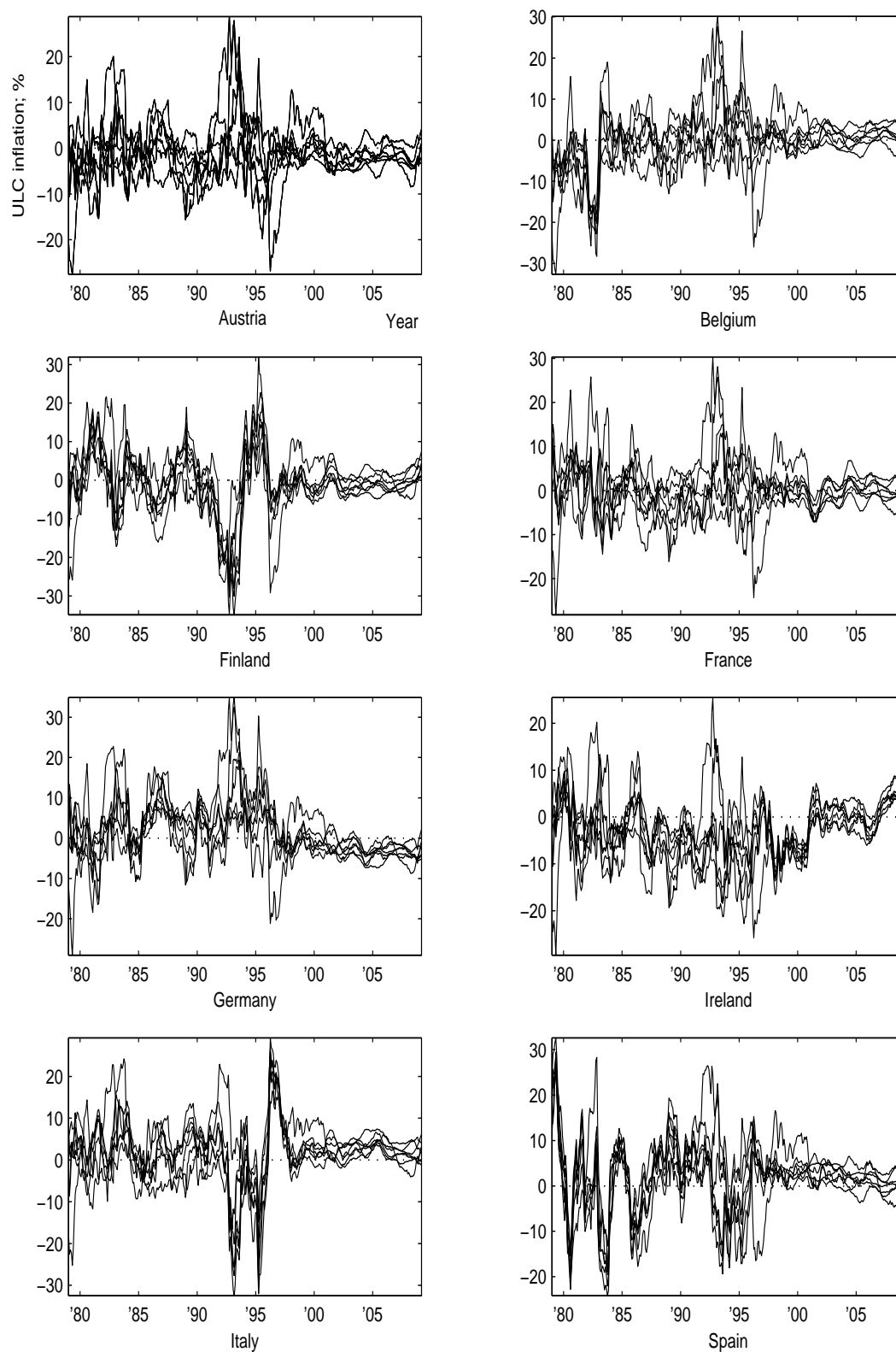


ULC growth differentials computed against each of the sample economies while Figure 7.4 displays the economy specific deviations from the cross sectional mean.

Each panel of Figure 7.3 reconfirms the marked reduction of volatility observed in Figure 7.1. Moreover, a high degree of cross sectional co-movement is apparent from the graphs. Thorough inspection of the graphs allows to identify some economies characterized by prolonged periods of above or below average ULC growth. The latter issue is highlighted by Figure 7.4. For the first subperiod ending with the introduction of the Euro, eyeball inspection reveals persistent positive deviations from the zero line for Germany while Ireland is characterized by a sustained improvement of its competitive position through below average ULC growth. In contrast, the recent subsample starting in 1999 shows persistent negative ULC growth differentials for Germany and Austria, while Spain, Italy and Belgium display continued above average ULC growth. These visual

7.5 Persistent inflation differentials and stability of the EMU

Figure 7.3: Country specific panels of ULC growth differentials



7.5 Persistent inflation differentials and stability of the EMU

observations are condensed in Table 7.3, documenting descriptive statistics of the data.

Table 7.3: Descriptive statistics

Economy	01/1979-03/2009		01/1979-12/1998		01/1999-03/2009	
	$mean(\tilde{y}_t)$	$std(\tilde{y}_t)$	$mean(\tilde{y}_t)$	$std(\tilde{y}_t)$	$mean(\tilde{y}_t)$	$std(\tilde{y}_t)$
AUT	-1.039	2.626	-0.907	3.094	-1.295	1.271
BEL	-0.336	4.044	-0.906	4.838	0.777	0.881
FIN	-0.639	5.863	-0.708	7.163	-0.502	1.217
FRA	0.071	2.567	0.256	2.930	-0.289	1.593
GER	1.266	4.427	3.022	4.428	-2.162	1.348
IRE	-2.569	3.836	-3.566	3.778	-0.624	3.152
ITA	1.596	4.988	1.264	6.080	2.245	0.878
ESP	1.649	5.404	1.545	6.599	1.851	1.127

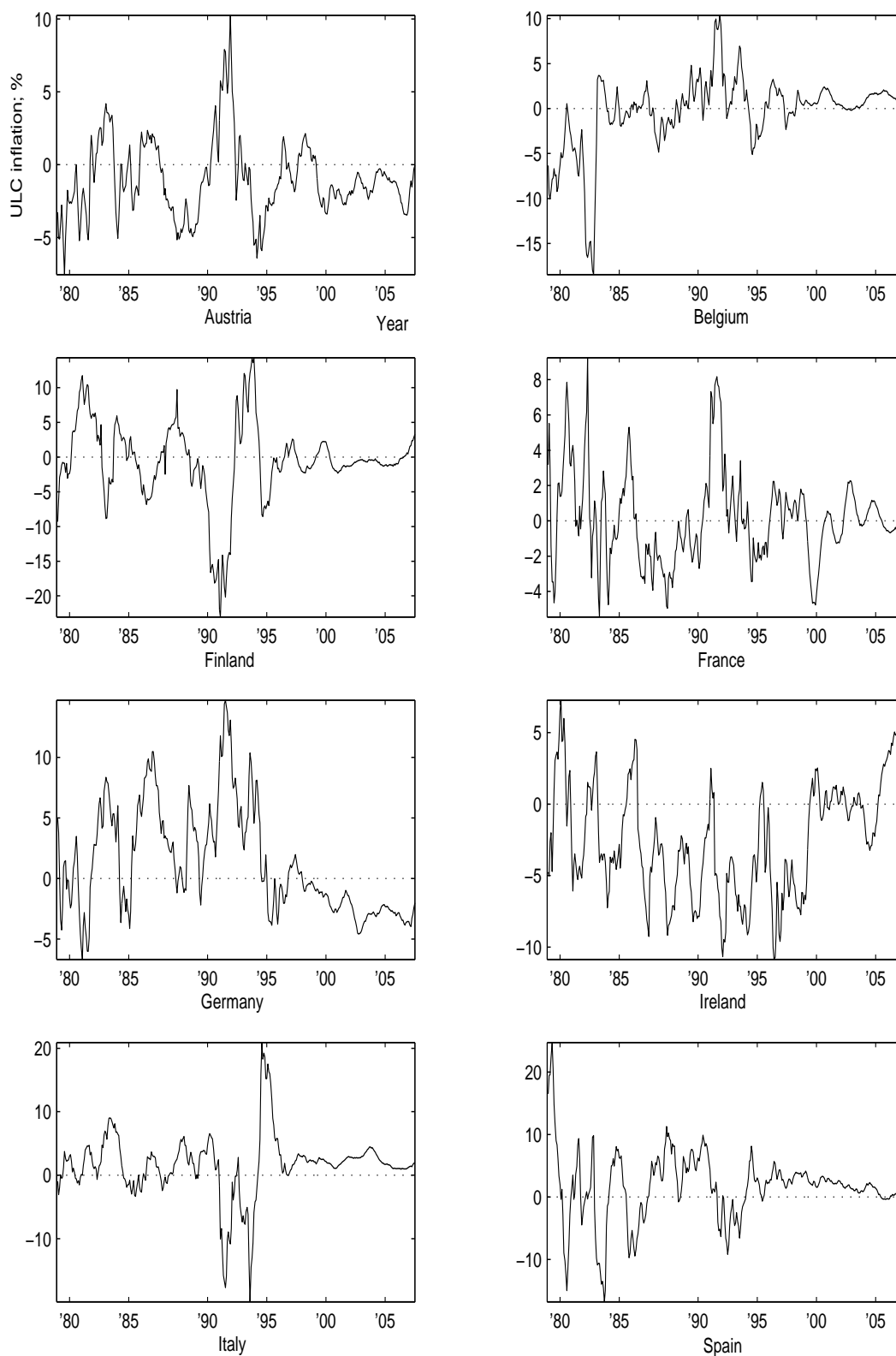
Notes: \tilde{y}_t denotes the ULC inflation differential computed with respect to the cross sectional mean.

7.5.3 Results

To assess the convergence hypothesis, PURTs are run for three subsamples of the data. Firstly, the entire sample period is considered. The results from this exercise are to be compared with the results of Dullien and Fritsche (2008), Fritsche and Kuzin (2007) and Fischer (2007) who (due to the low frequency of their data) only consider the whole time period. Subsequently, convergence tests are applied to the subsamples prior and after the Euro introduction. This approach is common in the literature of testing CPI inflation convergence in the EMU (e.g. Mentz and Sebastian, 2006, Buseti et al., 2007, or Lopez and Papell, 2008). Since it is the aim of our study to test absolute β -convergence, all tests are run without deterministic terms. Residual serial correlation is removed by means of the prewhitening scheme of Breitung and Das (2005) discussed in Chapter 4.3.2 and the cross sectional unit specific lag length is determined by means of the MAIC criterion of Ng and Perron (2001). Since the introduction of

7.5 Persistent inflation differentials and stability of the EMU

Figure 7.4: Country specific ULC growth differentials



7.5 Persistent inflation differentials and stability of the EMU

the Euro likely coincides with a structural change in the dynamics of the data, both subsamples are separately prewhitened.

The results of the PURTs are documented in Table 7.4. Results for the tests of group-wide convergence are reported at the top of the table and the results for country specific panels follow in alphabetical order. Considering the EU8 panel, the null hypothesis is rejected for the whole period as well as for the first subsample unanimously by all considered PURTs, irrespective if asymptotic or bootstrap critical values are used. For the last subsample, t -ratios are somewhat smaller in absolute values for all tests. Most noticeable, marginal significance levels obtained for the t_{Wh} statistic increase to 3.5% based on Gaussian critical values, respectively 5.0% based on bootstrap critical values. However, bearing in mind that the bootstrap variant t_{OLS}^* yields significantly higher power than t_{Rob}^* or t_{Wh}^* in finite samples, the hypothesis of group-wise convergence still holds even for the post Euro introduction subsample. This evidence in favor of overall convergence is in line with the PURT results presented in Dullien and Fritsche (2008), Fritsche and Kuzin (2007) and Fischer (2007). However, this finding has to be confronted with country specific evidence. Results for country specific panels of ULC differentials computed vis-à-vis the remaining sample economies are more heterogenous. While convergence is accepted for each of the considered economies for the entire sample period as well as for the first subperiod, analyses for the second subsample identify a number of panels where the null hypothesis of divergence cannot be rejected. Most clearly, divergent ULC growth behavior cannot be rejected for Germany and Spain with bootstrap p -values in excess of 20%, respectively 15%. It is noteworthy that bootstrap p -values are much larger than p -values based on the asymptotic approximation. This fact illustrates that the small cross sectional dimension, likely strong contemporaneous correlation as well as nonstationary volatility render the asymptotic approximations unreliable. The finding of divergence for Germany and Spain coincides with visual evidence and descriptive statistics presented above and illustrates a substantial improve-

7.5 Persistent inflation differentials and stability of the EMU

Table 7.4: Convergence of ULC growth differentials: PURT evidence

Sample	<i>T</i>	01/1979-03/2009			01/1979-12/1998			01/1999-03/2009				
		<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>T</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>	<i>T</i>	<i>OLS</i>	<i>rob</i>	<i>Wh</i>
EUS	346	-6.68 (.000) [.000]	-5.74 (.000) [.000]	-4.10 (.000) [.000]	225	-5.72 (.000) [.000]	-4.88 (.000) [.000]	-3.98 (.000) [.000]	110	-4.16 (.000) [.005]	-3.06 (.001) [.007]	-1.81 (.035) [.050]
AUT	346	-6.66 (.000) [.000]	-5.59 (.000) [.000]	-4.17 (.000) [.000]	225	-5.74 (.000) [.000]	-4.82 (.000) [.000]	-4.04 (.000) [.000]	110	-3.25 (.001) [.017]	-2.14 (.016) [.038]	-1.79 (.037) [.057]
BEL	346	-6.71 (.000) [.000]	-4.74 (.000) [.000]	-2.70 (.004) [.004]	225	-5.78 (.000) [.000]	-4.01 (.000) [.000]	-2.81 (.002) [.003]	110	-3.65 (.000) [.009]	-2.88 (.002) [.008]	-2.09 (.019) [.021]
FIN	349	-8.12 (.000) [.000]	-4.24 (.000) [.000]	-2.93 (.002) [.001]	226	-6.44 (.000) [.000]	-3.31 (.001) [.001]	-2.82 (.002) [.002]	110	-3.55 (.000) [.006]	-3.01 (.001) [.003]	-2.22 (.013) [.016]
FRA	346	-6.67 (.000) [.000]	-5.70 (.000) [.000]	-3.37 (.000) [.000]	225	-5.69 (.000) [.000]	-4.88 (.000) [.000]	-3.26 (.000) [.000]	110	-3.63 (.000) [.010]	-2.66 (.004) [.014]	-1.70 (.044) [.074]
GER	346	-5.57 (.000) [.000]	-3.50 (.000) [.001]	-2.67 (.004) [.004]	226	-4.87 (.000) [.000]	-3.11 (.001) [.001]	-2.71 (.003) [.003]	110	-1.61 (.087) [.250]	-1.06 (.186) [.289]	-0.97 (.201) [.313]
IRE	347	-5.27 (.000) [.000]	-3.31 (.001) [.001]	-3.10 (.001) [.001]	225	-3.94 (.000) [.001]	-2.60 (.005) [.007]	-2.71 (.003) [.006]	110	-6.05 (.000) [.018]	-2.78 (.003) [.024]	-1.62 (.053) [.085]
ITA	347	-6.80 (.000) [.000]	-3.87 (.000) [.001]	-2.63 (.004) [.004]	226	-5.80 (.000) [.000]	-3.35 (.000) [.001]	-2.62 (.005) [.004]	110	-2.76 (.003) [.046]	-2.24 (.013) [.037]	-1.66 (.049) [.078]
ESP	346	-7.17 (.000) [.000]	-4.14 (.000) [.000]	-3.34 (.000) [.000]	225	-6.69 (.000) [.000]	-3.89 (.000) [.000]	-3.53 (.000) [.000]	110	-2.67 (.004) [.177]	-1.61 (.053) [.152]	-1.23 (.110) [.206]

Notes: Reported numbers are *t*-ratios with associated Gaussian *p*-values in parentheses and numbers in square brackets are bootstrap *p*-values. Entries underneath *T* indicate the available number of observations after adjusting for the initial observations used in the prewhitening scheme. Unit specific lag lengths for the prewhitening are determined by means of the MAIC.

7.5 Persistent inflation differentials and stability of the EMU

ment (decline) of the relative competitive position of Germany (Spain) against the other considered EMU member economies, even when accounting for differences in productivity growth. For the remaining economies, bootstrap p -values are noticeable larger than in the first subsample. For Italy, Ireland, Austria, and France, convergence no longer holds at the 1% but only at the 5% (t_{OLS}^* , t_{Rob}^*) respectively 10% (t_{Wh}^*) significance level. The finding of growing dispersion of intra EMU competitiveness based on a productivity adjusted inflation measure is in contrast to results of Fischer (2007). He claims that while divergent CPI inflation dynamics can be identified after the start of the EMU, convergence still prevails if a productivity based measure is applied.

To corroborate these findings, additional univariate unit root tests are run on country specific deviations from the cross sectional mean. The set of M -unit root tests discussed in Ng and Perron (2001) (compare Chapter 2.4.2 and 2.5.1) and their wild bootstrap counterparts proposed by Cavaliere and Taylor (2008) are applied. Results are reported in Table 7.5. Considering the overall evidence, the univariate results indicate nonstationary for more cross sectional entities than the PURTs. For the first subperiod covering the time period until the introduction of the Euro, H_0 cannot be rejected for Ireland. Moreover, bootstrap p -values for Italy, Germany, Belgium and Spain increase beyond the 1% cutoff level. However, considering that H_0 is rejected for seven out of 8 sample economies at least at the 5% level, the results are supportive of convergent CPI inflation prior to the advent of the Euro and thus in line with the findings in the literature (Engel and Rogers, 2004, Beck and Weber, 2005 or Buseti et al., 2007). Compared to the first subsample, the number of convergent economies is substantially reduced for the time period following the introduction of the Euro. At the 1% significance level, convergence only holds for Finland and France, while a rejection of H_0 at the 5% level can be obtained for Ireland. For all other sample economies bootstrap p -values range between 0.146 and 0.676 with highest p -values obtained for Germany, Spain and Italy. However, it remains a-priori unclear if the increased

7.5 Persistent inflation differentials and stability of the EMU

Table 7.5: Convergence of ULC inflation differentials, univariate tests

Sample	T	01/1979-03/2009			T	01/1979-12/1998			T	01/1999-03/2009		
		MZ_a	MSB	MZ_t		MZ_a	MSB	MZ_t		MZ_a	MSB	MZ_t
AUT	349	-24.35 (.000) [.003]	0.14 (.000) [.003]	-3.49 (.000) [.003]	226	-19.63 (.004) [.002]	0.159 (.002) [.003]	-3.13 (.013) [.002]	110	-3.74 (.258) [.172]	0.37 (.243) [.237]	-1.37 (.360) [.146]
BEL	350	-16.01 (.009) [.014]	0.18 (.006) [.014]	-2.83 (.027) [.013]	227	-10.82 (.036) [.024]	0.22 (.028) [.028]	-2.33 (.076) [.023]	110	-2.91 (.334) [.235]	0.41 (.320) [.280]	-1.21 (.442) [.212]
FIN	350	-161.05 (.000) [.000]	0.06 (.000) [.000]	-8.97 (.000) [.000]	227	-114.43 (.000) [.000]	0.066 (.000) [.000]	-7.56 (.000) [.000]	121	-23.81 (.001) [.000]	0.14 (.000) [.000]	-3.23 (.010) [.001]
FRA	347	-19.05 (.004) [.004]	0.16 (.002) [.004]	-2.99 (.019) [.005]	226	-19.86 (.004) [.001]	0.16 (.002) [.001]	-3.15 (.012) [.001]	110	-52.74 (.000) [.000]	0.09 (.000) [.000]	-4.97 (.000) [.000]
GER	346	-11.19 (.033) [.028]	0.21 (.024) [.029]	-2.36 (.072) [.026]	226	-10.18 (.043) [.019]	0.22 (.033) [.023]	-2.25 (.088) [.016]	110	-0.55 (.745) [.536]	0.76 (.723) [.676]	-0.42 (.723) [.519]
IRE	346	-3.54 (.275) [.180]	0.37 (.257) [.215]	-1.32 (.382) [.161]	225	-1.60 (.522) [.373]	0.46 (.388) [.396]	-0.74 (.641) [.392]	110	-10.56 (.039) [.029]	0.22 (.030) [.039]	-2.29 (.081) [.026]
ITA	349	-15.13 (.012) [.031]	0.18 (.008) [.032]	-2.74 (.033) [.030]	226	-12.38 (.024) [.034]	0.20 (.017) [.039]	-2.48 (.056) [.029]	122	-0.65 (.712) [.501]	0.71 (.683) [.604]	-0.46 (.726) [.488]
ESP	347	-21.24 (.002) [.004]	0.15 (.001) [.004]	-3.25 (.009) [.004]	228	-12.07 (.026) [.018]	0.20 (.019) [.020]	-2.45 (.060) [.017]	110	-0.67 (.721) [.507]	0.65 (.619) [.558]	-0.43 (.718) [.505]

Notes: See Table 7.4.

number of divergent economies indicated by univariate tests is due to comparably lower power of univariate tests or to the power of PURTs against a heterogenous alternative. Moreover, as demonstrated in Banerjee et al. (2005), cross unit cointegration among the sample economies might render the considered PURTs oversized and hence bias panel evidence in favor of convergence. Accordingly, joint non-rejections of H_0 via PURTs and univariate tests are interpreted as robust evidence in favor of divergence. Similarly, joint rejections hint at robust evidence in favor of convergence. On the other hand, some ambiguity remains with respect to Austria, Belgium and Ireland for which conflicting results are obtained from panel and univariate tests. In the case of Ireland, a unit root

is detected for the first subsample by univariate tests, while panel evidence is rather in favor of stationarity. For the second subsample, a rejection of H_0 at the 5% significance level is obtained via both univariate and most panel tests. Combined results of panel and univariate tests for Austria and Belgium indicate convergent ULC growth prior to the introduction of the Euro. For the recent subsample, however, divergent behavior can be detected by univariate tests while bootstrap p -values of PURTs range between 1.7-5.7% and 0.8-2.1% for Austria and Belgium, respectively, thus rather indicating stationarity.

Considering the overall evidence, the introduction of the Euro helped to stabilize inflationary expectations at a low level as the volatility and cross sectional dispersion of ULC inflation has been markedly reduced. On the other hand, growing disparities of intra EMU competitive positions evolving after the introduction of the Euro can be detected by means of robust bootstrap PURTs. This result confirms the finding of divergent CPI inflation dynamics by Buseti et al. (2007) and points out that divergence among EMU economies cannot be (fully) attributed to the Balassa-Samuelson effect. While Germany has been improving its competitive position since 1999 (or reversing the deterioration of its competitiveness prior to the advent of the Euro), Spain and Italy have been persistently losing competitiveness against their trading partners within the EMU. On the other hand, wage setting in Finland and France was in line with stabilizing ULC growth at the cross sectional mean.

7.6 Conclusions

Convergence of ULC inflation among EMU economies is considered a requirement for the long run stability of the monetary union. In this paper, it is argued that prominent features of empirical data sets on ULC growth differentials call for the application of robust PURTs in testing the convergence hypothesis. In particular, a small cross sectional dimension, strong cross sectional correlation and marked

variance breaks lead to size distortions of existing homogenous PURTs. The asymptotic validity and consistency of wild bootstrap PURTs under such scenarios are established. Their satisfactory finite sample properties are demonstrated by means of a simulation study. In particular, it turns out that the bootstrap variant of the first generation PURT based on the OLS t -ratio of a pooled DF regression provides not only most reliable size control but also allows the researcher to make use of higher power in finite samples compared with the considered second generation tests. Subsequently, the issue of ULC growth convergence is investigated by testing for unit roots in ULC growth differentials. The data set consists of monthly data for eight EMU member economies and covers the period from January 1978 through March 2009. The empirical results highlight the importance of applying robust PURTs as bootstrap p -values are generally larger than p -values based on asymptotic approximations. While the period prior to the introduction of the Euro is characterized by converging ULC inflation rates, ULC inflation convergence is less prevalent in the recent period starting with the introduction of the Euro. Even though the volatility of ULC growth differentials has significantly dropped as a consequence of the Euro introduction, there are some economies which have witnessed marked shifts in their relative price competitiveness compared with their EMU trading partners. By combining evidence from panel and univariate tests, those economies with diverging ULC growth are identified. Most notably, Germany has been improving its competitive position, while above average ULC growth in Italy and Spain has impaired their relative price competitiveness.

While the connection between accumulating real exchange rate changes and growing trade imbalances within the EMU is immediate, further adverse long term effects may arise from diverging inflationary dynamics not matched by equal productivity growth. First, as pointed out by Dullien and Fritsche (2008), economies which display inflation rates persistently above the EMU average are more likely to suffer from investment misallocations. For instance, the likelihood

of bubbles in the housing sector increases: Within a monetary union, above average inflation makes finance less costly while, at the same time, the deteriorating external competitiveness reduces the attractiveness of investments in the tradeables sector. As an illustrative example, consider the case of Spain. Over the last decade, Spain has been characterized by sustained above average ULC and CPI inflation rates. As a consequence of the implied deterioration of its external competitiveness, the Spanish current account deficit has been successively widening and amounted to over 9% of GDP in the past two years. Internally, Spain witnessed a tremendous increase in house prices with an associated boom in construction. In the course of the current financial crisis, Spain has been among the most affected economies in the Euro area. The bursting of the housing bubble and the end of the construction boom has led to surging unemployment. According to Eurostat, the unemployment rate averaged 8.3% in 2007 and has more than doubled since then. Even though the illustrative case study of Spain is not firmly backed by econometric analyses, it points out that persistently diverging ULC developments within the EMU may be indicative of more serious economic imbalances.

Chapter 8

Conclusions

The first part of the thesis is a disquisition on univariate unit root tests. Testing for unit roots in univariate time series models not only constitutes a huge literature in itself but also provides the basis for the panel unit root tests (PURT) discussed in later Chapters. Therefore, a review over the relevant literature is provided in Chapter 2. Treatment of serial correlation as well as the development of more powerful test procedures are the focal point in the evolution of this literature. Nonparametric or simulation and resampling based tests are also discussed since these concepts are reconsidered in later chapters. In particular, a simulation based approach to unit root testing is developed in Chapter 3. The underlying principle, namely to regress the data on computer generated random walks, makes use of an idea suggested by Park (1990). Even though formal proofs of the new test's validity and consistency have yet to be established, the finite sample simulation results indicate that the discussed approach is a viable alternative to test the unit root hypothesis which has not been considered before. Furthermore, the reported finite sample results are rather promising, encouraging further research on this approach. In particular, given its nonparametric foundations, it is interesting to see how this test performs with respect to outliers or nonlinear data transformations.

The second part of this work concentrates on homogenous PURTs. While a review over the literature on homogenous PURTs is provided in Chapter 4, extensions of existing test procedures are developed in Chapters 5, 6, and 7. In Chapter 5, a ‘White’-corrected PURT statistic is proposed which is robust under weak form cross sectional dependence. For models with strong form cross sectional dependence, wild bootstrap PURTs are suggested. The finite sample characteristics of the ‘White-type’ statistic are found to be similar as a competing second generation PURT statistic of Breitung and Das (2005). Irrespective of the assumed contemporaneous correlation structure, best finite sample results are offered by the wild bootstrap tests. The characteristics of the ‘White-type’ statistic proposed in Chapter 5 are then further analyzed in Chapter 6. While the adverse effects of volatility breaks on univariate unit root tests are already documented in the literature, this is not the case for PURTs. In Chapter 6, it is proven the proposed statistic remains asymptotically pivotal under discrete breaks in the innovation variance. In contrast, established homogenous PURTs are shown to depend asymptotically on nuisance parameters in this case. By further relaxing the assumptions on the innovations of the data generating process, the performance of the proposed wild bootstrap PURTs are assessed in Chapter 7. The validity of the wild bootstrap PURTs introduced in Chapter 5 is validated even under very general cases of cross sectional dependence and nonstationary heteroskedasticity. A simulation study confirms these theoretical findings, illustrating the robust size control of all bootstrap PURTs in finite samples. Moreover, it is shown that a the bootstrap implementation of a first generation PURT may offer power gains compared with competing second generation tests.

The economic applications in this thesis not only illustrate the functioning of the discussed tests in applied work. Moreover, possible fallacies of application of tests which impose too restrictive assumptions on the data generating process are highlighted. The univariate test results on the purchasing power parity (PPP) hypothesis in Chapter 3 indicate that weak form PPP only prevails for a minority

of the sample economies when the US are taken as the benchmark economy. This finding is in line with similar univariate analyses in the literature. One interesting aspect of the analysis in Chapter 3 is that the proposed test yields results which are in accordance with the more powerful of the established tests. In contrast, the traditional (augmented) Dickey-Fuller and the nonparametric test of Park (1990) detect even less instances of stationary real exchange rates. This result reflects the fact that low power is an issue of great empirical concern in univariate unit root testing. The empirical example in Chapter 5 investigates the question if, despite widely observed persisting current account imbalances, intertemporal budget restrictions are binding. Therefore, the order of integration of current account balances for up to 129 economies is tested by means of different PURTs. The resulting evidence strongly suggests that current account balances are in line with intertemporal budget constraints, at least for a non-zero fraction of the considered economies. By complementing the PURT findings by a summary of individual specific univariate tests, the power deficiency of the latter are underscored once more. Moreover, significantly smaller t -ratios obtained via a first generation PURT compared with second generation tests implicitly point out the size distortions associated with neglecting cross sectional dependence when testing for unit roots in panel data sets. The importance of accounting not only for cross sectional dependence but also for potential volatility shifts in the data is demonstrated in the economic application of Chapter 6. While evidence in favor of the Fisher hypothesis (implying stationary real interest rates) is supported by all considered PURTs, differences among the tests prevail when testing the order of integration of the inflation and nominal interest rates. In these instances, the volatility robust test yields notable larger marginal significance levels compared with the included benchmark tests. In Chapter 7, convergence of unit labor costs in the European Economic and Monetary Union (EMU) is tested by means of panel and univariate unit root tests. Relying on the results obtained via robust wild bootstrap implementations of previously discussed test procedures, some

evidence of a widening of competitive positions within the EMU after the introduction of the Euro can be found. As before, it is demonstrated that time varying volatility is a central stylized fact of many macroeconomic time series.

References

- Acosta, C. P. (2007). Inflation differentials and business cycle fluctuations in the European Monetary Union. IMK Working Paper 5/2007.
- Aparico, F., Escribano, A., and Siplis, A. E. (2006). Range unit-root (rur) tests: robust against nonlinearities, error distributions, structural breaks and outliers. *Journal of Time Series Analysis*, 27(4):545–576.
- Bai, J. and Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, 72(4):1127–1178.
- Balassa, B. (1964). The purchasing-power parity doctrine: a reappraisal. *The Journal of Political Economy*, 72(6):584–596.
- Baltagi, B. H., Bresson, G., and Pirotte, A. (2007). Panel unit root tests and spatial dependence. *Journal of Applied Econometrics*, 22(2):339–360.
- Baltagi, B. H. and Kao, C. D. (2000). Nonstationary panels, cointegration in panels and dynamic panels: a survey. In Baltagi, B., editor, *Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, pages 7–52. JAI, Amsterdam.
- Banerjee, A., Marcellino, M., and Osbat, C. (2005). Testing for PPP: Should we use panel methods? *Empirical Economics*, 30(1):77–91.
- Barro, R. and Sala-i-Martin, X. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, 1991(1):107–182.

REFERENCES

- Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H., and Taylor, R. L. (1991). Bootstrapping unstable first-order autoregressive processes. *Annals of Statistics*, 19(2):1098–1101.
- Beck, G. and Weber, A. A. (2005). Inflation rate dispersion and convergence in monetary and economic unions: Lessons for the ECB. CFS Working Paper 2005/31.
- Beck, N. and Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review*, 89(3):634–647.
- Bernard, A. B. and Durlauf, S. N. (1996). Interpreting tests of the convergence hypothesis. *Journal of Econometrics*, 71(1):161–173.
- Bhargava, A. (1986). On the theory of testing for unit roots in observed time series. *Review of Economic Studies*, 53(3):369–384.
- Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- Breitung, J. (2000). The local power of some unit root tests for panel data. In Baltagi, B., editor, *Advances in Econometrics, Vol. 15*, pages 161–178. JAI, Amsterdam.
- Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. *Journal of Econometrics*, 108(2):343–363.
- Breitung, J. and Das, S. (2005). Panel unit root tests under cross sectional dependence. *Statistica Neerlandica*, 59(4):414–433.
- Breitung, J. and Das, S. (2008). Testing for unit roots in panels with a factor structure. *Econometric Theory*, 24(1):88–108.
- Breitung, J. and Gouriéroux, C. (1997). Rank tests for unit roots. *Journal of Econometrics*, 81(1):7–27.

REFERENCES

- Breitung, J. and Meyer, W. (1994). Testing for unit roots in panel data: are wages on different bargaining levels cointegrated? *Applied Economics*, 26(4):353–361.
- Breitung, J. and Pesaran, H. M. (2008). Unit roots and cointegration in panels. In Matyas, L. and Sevestre, P., editors, *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Kluwer Academic Publishers, Dordrecht.
- Breitung, J. and Westerlund, J. (2009). Myths and facts about panel unit root tests. Göteborg University, Department of Economics Working Paper 380.
- Busetti, F., Forni, L., Harvey, A., and Venditti, F. (2007). Inflation convergence and divergence within the European Monetary Union. *International Journal of Central Banking*, 3(2):95–121.
- Caporale, G. M. and Pittis, N. (1993). Common stochastic trends and inflation convergence in the EMS. *Review of World Economics*, 129(2):207–215.
- Cavaliere, G. (2004). Unit root tests under time varying variances. *Econometric Reviews*, 23(4):259–292.
- Cavaliere, G., Rahbek, A., and Taylor, A. M. R. (2010). Testing for co-integration in vector autoregressions with non-stationary volatility. *Forthcoming in: Journal of Econometrics*.
- Cavaliere, G. and Taylor, A. M. R. (2007a). Testing for unit roots in autoregressions with multiple level shifts. *Econometric Theory*, 23(6):1162–1215.
- Cavaliere, G. and Taylor, A. M. R. (2007b). Testing for unit roots in time series with non-stationary volatility. *Journal of Econometrics*, 140(2):919–947.
- Cavaliere, G. and Taylor, A. M. R. (2008). Bootstrap unit root tests for time series with nonstationary volatility. *Econometric Theory*, 24(1):43–71.

REFERENCES

- Cavaliere, G. and Taylor, A. M. R. (2009). Bootstrap M unit root tests. *Econometric Reviews*, 28(5):393–421.
- Chang, Y. (2002). Nonlinear IV unit root test in panels with cross sectional dependency. *Journal of Econometrics*, 110(2):261–292.
- Chang, Y. (2004). Bootstrap unit root test in panels with cross sectional dependency. *Journal of Econometrics*, 120(2):263–294.
- Chang, Y. and Park, J. Y. (2003). A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis*, 24(4):379–400.
- Chang, Y. and Song, W. (2002). Panel unit root tests in the presence of cross-sectional dependency and heterogeneity. Mimeo.
- Chang, Y. and Song, W. (2005). Unit root tests for panels in the presence of short-run and long-run dependencies: Nonlinear IV approach with fixed N and large T . Rice University, unpublished.
- Cheung, Y.-W. and Lai, K. S. (1995). Lag order and critical values of the Augmented Dickey-Fuller test. *Journal of Business and Economic Statistics*, 13(3):277–280.
- Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, 20(2):261–292.
- Choi, I. (2006). Combination unit root tests for cross-sectionally correlated panels. In Corbae, D., Durlauf, S., and Hansen, B., editors, *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*. Cambridge University Press, New York.
- Choi, I. and Chue, T. K. (2007). Subsampling hypothesis for nonstationary panels with applications to exchange rates and stock prices. *Journal of Applied Econometrics*, 22(2):233–264.

REFERENCES

- Coakley, J., Kulasi, F., and Smith, R. (1998). The Feldstein Horioka puzzle and capital mobility - A review. *International Journal of Finance and Economics*, 3(2):169–188.
- Cooray, A. (2003). The Fisher effect: a survey. *The Singapore Economic Review*, 48(2):135–150.
- Crowder, W. J. (2003). Panel estimates of the Fisher effect. *University of Texas at Arlington, Discussion Paper*.
- Davidson, R. and Flachaire, E. (2001). The wild bootstrap, tamed at last. *GRE-QAM Document de Travail 99A32*.
- DeJong, D. N., Nankervis, J. C., Savin, N. E., and Whiteman, C. H. (1992). The power problems of unit root tests in time series with autoregressive errors. *Journal of Econometrics*, 53(3):323–343.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(386):427–431.
- Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio test for autoregressive time series with a unit root. *Econometrica*, 49(4):1051–1072.
- Donsker, M. D. (1951). An invariance principle for certain probability limit theorems. *Memoirs of the American Mathematical Society*, 6:1–12.
- Dullien, S. and Fritsche, U. (2008). Does the dispersion of unit labor cost dynamics in the EMU imply long-run divergence? Results from a comparison with the United States of America and Germany. *International Economics and Economic Policy*, 5(3):269–295.
- ECB (2005). Monetary policy and inflation differentials in a heterogeneous currency area. *ECB Monthly Bulletin*, May:61–78.

REFERENCES

- Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26(3):244–268.
- Elliott, G. (1999). Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review*, 40(3):767–783.
- Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836.
- Engel, C. and Rogers, J. H. (2004). European product market integration after the euro. *Economic Policy*, 19(39):347–384.
- Feldstein, M. and Horioka, C. (1980). Domestic savings and international capital flows. *Economic Journal*, 90(June):314–329.
- Fischer, C. (2007). An assessment of the trends in international price competitiveness among EMU countries. Deutsche Bundesbank Discussion Paper, Series 1: Economic Studies 8/2007.
- Fisher, I. (1930). *The Theory of Interest*. Macmillan, New York.
- Fisher, R. A. (1954). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.
- Fritsche, U. and Kuzin, V. (2007). Unit labor cost growth differentials in the Euro area, Germany, and the US: lessons from PANIC and cluster analysis. DEP Discussion Papers, Macroeconomics and Finance Series, 3/2007.
- Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2(2):111–120.
- Greene, W. H. (2003). *Econometric Analysis*. Prentice Hall, New Jersey.

REFERENCES

- Hamori, S. and Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift. *Economics Letters*, 57(3):245–253.
- Hanck, C. (2009a). Cross sectional correlation robust tests for panel cointegration. *Journal of Applied Statistics*, 37(1):817–833.
- Hanck, C. (2009b). Nonstationary-volatility robust panel unit root tests and the great moderation. *Mimeo*.
- Harvey, A. and Bates, D. (2003). Multivariate unit root tests, stability and convergence. University of Cambridge, DAE Working Paper No. 301.
- Herwartz, H. (2006). Testing for random effects in panel data under cross sectional error correlation - A bootstrap approach to the Breusch Pagan test. *Computational Statistics and Data Analysis*, 50(12):3567–3591.
- Herwartz, H. and Neumann, M. H. (2005). Bootstrap inference in systems of single equation error correction models. *Journal of Econometrics*, 128(1):165–193.
- Herwartz, H. and Reimers, H. E. (2006). Panel nonstationarity tests of the Fisher hypothesis: an analysis of 114 economies during the period 1960-2004. *Applied Econometrics and International Development*, 6(3):39–54.
- Herwartz, H. and Reimers, H. E. (2009). A functional coefficient approach to modeling the fisher hypothesis: World wide evidence. *forthcoming in: Macroeconomic Dynamics*.
- Heston, A., Summers, R., and Aten, B. (2006). Penn world tables version 6.2. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Horowitz, J. L. (2001). The bootstrap. In Heckman, J. J. and Leamer, E., editors, *Handbook of Econometrics*. Elsevier, Amsterdam.

REFERENCES

- Hurlin, C. and Mignon, V. (2007). Second generation panel unit root tests. *Manuscript, THEMA-CNRS, University of Paris X.*
- Im, K. S., Pesaran, H. M., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.
- Jönsson, K. (2005). Cross-sectional dependency and size distortion in a small-sample homogeneous panel data unit root test. *Oxford Bulletin of Economics and Statistics*, 67(3):369–392.
- Kaliva, K. (2008). The fisher effect, survey data and time-varying volatility. *Empirical Economics*, 35(1):1–10.
- Kim, C. J. and Nelson, C. R. (1999). Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics*, 81(4):608–616.
- Kim, T. H., Leybourne, S., and Newbold, P. (2002). Unit root tests with a break in innovation variance. *Journal of Econometrics*, 109(2):365–387.
- Levin, A., Lin, C. F., and Chu, C. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1):1–24.
- Liu, R. Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Annals of Statistics*, 16(4):1696–1708.
- Lopez, C. and Papell, D. H. (2008). Testing for group-wise convergence with an application to Euro area inflation. University of Cincinnati Working Paper Series.
- MacKinnon, J. G. and White, H. (1985). Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3):305–329.

REFERENCES

- Maddala, G. S. and Wu, S. (1999). A comparative study of unit root tests with panel data and a simple new test. *Oxford Bulletin of Economics and Statistics*, Special Issue:631–652.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21(1):255–285.
- McConnell, M. M. Perez-Quiroz, G. (2000). Output fluctuations in the United States: what has changed since the early 1980s? *American Economic Review*, 90(5):1464–1476.
- McGarvey, M. G. and Walker, M. B. (2003). An alternative estimator for fixed effects spatial models. *Georgia State University, Working Paper 03-03*.
- Mentz, M. and Sebastian, S. P. (2006). Inflation convergence after the introduction of the Euro. In Columbus, F., editor, *Inflation: Researches and Perspectives*. Hauppauge, New York.
- Moon, H. R. and Perron, B. (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics*, 122(1):81–126.
- Moon, H. R. and Perron, B. (2008). Asymptotic local power of pooled t-ratio tests for unit roots in panels with fixed effects. *Econometrics Journal*, 11(1):80–104.
- Moon, H. R., Perron, B., and Phillips, P. C. B. (2006). On the Breitung test for panel unit roots and local asymptotic power. *Econometric Theory*, 22(6):1179–1190.
- Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.
- Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69(6):1519–1554.

REFERENCES

- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1399–1416.
- O’Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics*, 44(1):1–19.
- Palm, F. C., Smeekes, S., and Urbain, J.-P. (2008a). Bootstrap unit root tests. *Journal of Time Series Analysis*, 29(2):371–401.
- Palm, F. C., Smeekes, S., and Urbain, J.-P. (2008b). Cross sectional dependence robust block bootstrap panel unit root tests. Maastricht, METEOR Research Memoranda 048.
- Paparoditis, E. and Politis, D. N. (2003). Residual-based block bootstrap for unit root testing. *Econometrica*, 71(3):813–855.
- Paparoditis, E. and Politis, D. N. (2005). Bootstrapping unit root tests for autoregressive time series. *Journal of the American Statistical Association*, 100(2):543–553.
- Park, J. Y. (1990). Testing for unit roots and cointegration by variable addition. In Fromby, T. and Rhodes, F., editors, *Advances in Econometrics, Vol 8: Cointegration, Spurious Regression and Unit Roots*. JAI Press, Cambridge.
- Park, J. Y. (2002). An invariance principle for sieve bootstraps in time series. *Econometric Theory*, 18(2):469–490.
- Park, J. Y. (2003). Bootstrap unit root tests. *Econometrica*, 71(6):1845–1895.
- Park, J. Y. and Choi, B. (1988). A new approach to testing for a unit root. CAE Working Paper No. 88-23, Cornell University, Ithaca, NY.
- Park, J. Y., Ouliaris, S., and Choi, B. (1988). Spurious regression and tests for cointegration. CAE Working Paper No. 88-07, Cornell University, Ithaca, NY.

REFERENCES

- Perron, P. and Ng, S. (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies*, 63(3):435–463.
- Perron, P. and Ng, S. (1998). An autoregressive spectral density estimator at frequency zero for nonstationarity tests. *Econometric Theory*, 14(5):560–603.
- Perron, P. and Qu, Z. (2007). A simple modification to improve the finite sample properties of Ng and Perron’s unit root test. *Economics Letters*, 94(1):12–19.
- Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics*, 22(2):265–312.
- Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogenous panels. *Journal of Econometrics*, 68(1):79–113.
- Pesaran, M. H., Smith, R., and Im, K. S. (1995). Dynamic linear models for heterogenous panels. DAE Working Paper 9503. University of Cambridge, Cambridge (UK).
- Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33(3):331–340.
- Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica*, 55(2):277–301.
- Phillips, P. C. B. and Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Journal of Econometrics*, 67(5):1057–1111.
- Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.
- Phillips, P. C. B. and Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, 6(1):217–259.

REFERENCES

- Psaradakis, Z. (2001). Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors. *Journal of Time Series Analysis*, 22(5):577–594.
- Quah, D. (1994). Exploiting cross section variation for unit root inference in dynamic data. *Economic letters*, 44(1):9–19.
- Rapach, D. E. and Weber, C. E. (2004). Are real interest rates really non-stationary? new evidence from tests with good size and power. *Journal of Macroeconomics*, 26(3):409–430.
- Rose, A. K. (1988). Is the real interest rate stable? *The Journal of Finance*, 43(5):1095–1112.
- Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607.
- Samuelson, P. A. (1964). Theoretical notes on trade problems. *The Review of Economics and Statistics*, 46(2):145–154.
- Sargan, J. D. and Bhargava, A. (1983). Testing for residuals from least squares regression being generated by Gaussian random walk. *Econometrika*, 51(1):153–174.
- Stock, J. H. (1999). A class of tests for integration and cointegration. In Engle, R. F. and White, H., editors, *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger*. Oxford University Press, Oxford.
- Stock, J. H. and Watson, M. W. (2003). Has the business cycle changed and why? *NBER Macroeconomic Annual*, 17:159–230.
- Swensen, A. R. (2003). Bootstrapping unit root tests for integrated processes. *Journal of Time Series Analysis*, 24(1):99–126.

REFERENCES

- Taylor, A. M. and Taylor, M. P. (2004). The purchasing power parity debate. *Journal of Economic Perspectives*, 18(4):135–158.
- Wagner, M. (2008). On ppp, unit roots and panels. *Empirical Economics*, 35(2):229–249.
- White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, 48(4):817–838.
- White, H. (2000). *Asymptotic Theory for Econometricians*. Academic Press, San Diego.
- Wu, C. F. J. (1986). Jackknife bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, 14(4):1261–1295.
- Zellner, A. (1962). An efficient method for estimating seemingly unrelated regressions and tests of aggregation bias. *Journal of the American Statistical Association*, 57(298):348–368.

Lebenslauf

Persönliche Daten

Name	Florian Siedenburg
Geburtsdatum	8.10.1979
Geburtsort	Oldenburg
Staatsangehörigkeit	Deutsch
Familienstand	Verheiratet

Lebenslauf

05/1999	Abitur am Schulzentrum Geschwister-Scholl, Bremerhaven
09/1999 - 10/2000	Zivildienst
10/2000 - 08/2002	Studium der Volkswirtschaftslehre an der Christian-Albrechts-Universität zu Kiel
09/2002 - 05/2003	Studium an der Universidade Nova in Lissabon im Rahmen des Erasmus-Austauschprogramms
06/2003 - 01/2006	Studium der Volkswirtschaftslehre an der Christian-Albrechts-Universität zu Kiel, Schwerpunkt quantitative Wirtschaftsforschung Abschluss: Diplom-Volkwirt
01/2006-07/2010	Doktorandenstudium im Rahmen der Programms “Quantitative Economics” der Christian-Albrechts- Universität zu Kiel Abschluss: Promotion zum Dr. sc. pol.
05/2010 -	Angestellter bei der meteolytix GmbH, Kiel