

Heteroskedasticity-robust unit root testing for trending panels

Helmut Herwartz^a, Simone Maxand^b and Yabibal M. Walle^{*a}

^aDepartment of Economics, University of Goettingen, Germany

^bDepartment of Political and Economic Studies, University of Helsinki, Finland

Abstract Time-varying volatility and linear trends are common features of several macroeconomic time series. Recent papers have proposed panel unit root tests (PURT) that are pivotal in the presence of volatility shifts, excluding linear trends, however. This paper proposes a new PURT that works well for data that is both heteroskedastic and trending. Under the null hypothesis, the test statistic has a limiting Gaussian distribution. We derive the local asymptotic power to underpin the consistency of the test statistic. Simulation results reveal that the test performs well in small samples. As an empirical illustration, we examine the stationarity of energy use per capita in OECD economies. While the series are in general difference stationary, they could also be considered as trend stationary for specific time spans.

MOS codes: 60F05, 60F17, 62M10, 91B84.

JEL codes: C23, C12, Q40.

Keywords: Panel unit root tests, nonstationary volatility, cross-sectional dependence, near epoch dependence, energy use per capita.

*Corresponding author: Department of Economics, University of Goettingen, Humboldtallee 3, D-37073 Goettingen, Germany, tel.: +49 551 3921392, fax: +49 551 397279, e-mail: ywalle@uni-goettingen.de. We thank Jörg Breitung, Matei Demetrescu and two anonymous referees for helpful comments and suggestions. Simone Maxand gratefully acknowledges financial support from the Academy of Finland (308628).

1 Introduction

Many macroeconomic time series are characterized by two stylized features: time-varying (co)variances and persistent or transitory deviations from a linear trend. Regarding the former characteristic, e.g., Sensier and Dijk (2004) document that about 80% of 214 U.S. macroeconomic time series they studied displayed breaks in their unconditional volatility during the period 1959–1999. On the latter characteristic, Westerlund (2015, p. 454) states that

“...for many economic time series, a linear trend, rather than a constant, might be considered appropriate as the default specification, This is certainly true for series such as GDP, industrial production, money supply and consumer or commodity prices, where trending behavior is evident.”

The potential consequences of variance shifts on the discrimination between persistent and transitory dynamics by means of univariate unit root tests have been investigated by, among others, Hamori and Tokihisa (1997), Kim et al. (2002), Cavaliere (2004), and Cavaliere and Taylor (2007). These studies find that the (augmented) Dickey-Fuller (Dickey and Fuller, 1979) tests have seriously distorted empirical sizes—and, hence, provide deceptive inference—if volatility varies over time.

For the case of diagnosing unit roots in dynamic panels, the joint occurrence of time-varying volatility and linear trends complicates inference even further. First, widely applied PURTs such as those suggested in Levin et al. (2002) or Breitung and Das (2005) are no longer pivotal under heteroskedasticity (Herwartz et al., 2016). Second, even recently proposed heteroskedasticity-robust PURTs (Demetrescu and Hanck, 2012a,b; Herwartz et al., 2016) lack pivotalness under linear trends, since available detrending schemes introduce nuisance parameters that affect the limiting distribution of the tests under heteroskedasticity.

In this paper, we propose a new PURT that works well for the empirically-relevant case that the data is both trending and heteroskedastic. The test statistic builds upon the recursive detrending method proposed in Chang (2002). The drift term is estimated as the unconditional mean of first-differenced series. Tracing the effects of volatility breaks, level detrending and drift estimation, we construct a test statistic that exhibits an asymptotic Gaussian distribution under the panel unit root null hypothesis. To prove asymptotic normality we rely on central limit theory (CLT) for near-epoch dependent processes as discussed, e.g., in Davidson (1994). CLT of this type appears flexible and useful in practice, as it allows for a regulated dependence structure in the process under study. Results for the local asymptotic power of the test underpin its consistency. Simulation results show that the proposed test works well in finite samples, and

has satisfactory power which is comparable with the power of the tests in Demetrescu and Hanck (2012a) and Herwartz et al. (2016) under homoskedasticity.

As an empirical illustration, we examine whether energy use per capita is trend or difference stationary. Using data from 23 OECD economies over the period 1960–2014, we find that energy use per capita is generally integrated of order one. However, results from unit root testing for rolling fixed-length time spans show that the series could be characterized as trend stationary for forty-years windows that start between 1963 and 1968. It is worth mentioning here that the proposed test is implemented in the Stata package *xtpurt* (Herwartz et al., 2018).

Section 2 discusses the panel unit root testing problem and presents the proposed test statistic along with its asymptotic distribution. The finite sample performance of the new test is evaluated by means of a Monte Carlo study documented in Section 3. As an empirical illustration, the stationarity of energy use per capita is examined in Section 4. Section 5 concludes. Proofs of the asymptotic results are provided in the online Supporting Information.

2 Homogeneous panel unit root testing

In this section we, firstly, describe the panel unit root testing problem and formalize cross-sectional dependence and heteroskedasticity. Secondly, we recall the detrending scheme of Chang (2002), which is essential to establish pivotalness of the new test statistic. Thirdly, we present the new test in two versions: a theoretical (but infeasible) variant building upon true deterministic parameters, and a feasible approach based on estimated moments. Finally, we sketch the adjustment of the panel data in presence of higher order serial correlation.

2.1 The first order panel autoregression

The first order panel autoregression considered in this work formalizes an empirically-relevant panel unit root testing problem of distinguishing between a random walk with drift and a trend stationary process (Pesaran, 2007), i.e.,

$$\mathbf{y}_t = \boldsymbol{\mu} + (1 - \rho)\boldsymbol{\delta}t + \rho\mathbf{y}_{t-1} + \mathbf{e}_t, \quad t = 1, \dots, T. \quad (1)$$

In (1), $\mathbf{y}_t = (y_{1t}, \dots, y_{Nt})'$, $\mathbf{y}_{t-1} = (y_{1,t-1}, \dots, y_{N,t-1})'$, $\mathbf{e}_t = (e_{1t}, \dots, e_{Nt})'$ are $N \times 1$ vectors, and \mathbf{e}_t is heterogeneously distributed with mean zero and covariance Ω_t . Furthermore, the vector $\boldsymbol{\delta} = (\delta_1, \dots, \delta_N)'$ stacks panel-specific trend parameters, and $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)'$ contains panel-specific intercepts. PURTs are used to test the hypothesis $H_0 : \rho = 1$ against $H_1 : \rho < 1$ in (1).

While all PURTs share the same null hypothesis, they differ in their assumptions about the alternative hypothesis. The alternative hypothesis for PURTs such as the one suggested by Levin et al. (2002) is that the time series processes are stationary for all cross section members with *homogeneous* autoregressive parameter. Other PURTs, such as the one by Im et al. (2003), assume under the alternative hypothesis that a significant fraction of the panel members are stationary with *heterogeneous* (individual specific) mean reversion parameters. While homogeneous PURTs build upon pooled regressions, heterogeneous PURTs combine test statistics for individual panel units. However, it is worth noting that both groups of tests could be consistent against both homogeneous and heterogeneous alternatives. Moreover, rejection of the null hypothesis by both types of tests indicates that a significant fraction of the time series processes are stationary (Breitung and Pesaran, 2008; Pesaran, 2012 ; Westerlund and Breitung, 2013).

With strengthened moment conditions, we adopt the assumptions in Herwartz et al. (2016) about cross-sectional dependence and heteroskedasticity e_t as follows:

Assumptions \mathcal{A} .

- (i) e_t is serially uncorrelated with mean 0 and covariance Ω_t .
- (ii) Ω_t is a positive definite matrix with eigenvalues $\lambda_t^{(1)} \leq \lambda_t^{(2)} \leq \dots \leq \lambda_t^{(N)}$ and $\lambda_t^{(N)} < \bar{c} < \infty$, $\lambda_t^{(1)} > \underline{c} > 0$ for all t .
- (iii) $E[u_{it}^p u_{jt}^p u_{kt}^p u_{lt}^p] < \infty$ for all i, j, k, l and $p = 1, 2$, where $u_{\bullet t}$, $\bullet \in \{i, j, k, l\}$ denote typical elements of $\mathbf{u}_t = \Omega_t^{-1/2} e_t$. Here we set $\Omega_t^{1/2} = \Gamma_t \Lambda_t^{1/2} \Gamma_t'$, where Λ_t is a diagonal matrix of eigenvalues of Ω_t and the columns of Γ_t are the corresponding eigenvectors.

$\mathcal{A}(i)$ restricts the error terms to be serially uncorrelated. Ways of handling higher order serial correlation will be described later. The assumption that the fourth order moments of e_{it} (or u_{it} by implication of $\mathcal{A}(ii)$) should be finite ($\mathcal{A}(iii)$ for $p = 1$) is standard in the (panel) unit root literature. The stronger assumption of finiteness of moments up to order eight ($p = 2$) will allow to apply asymptotic theory for near-epoch dependent processes. While $\mathcal{A}(ii)$ captures so-called weak forms of cross-sectional dependence such as spatial panel models (for more details on spatial panel models see, e.g., Anselin, 2013) and seemingly unrelated regressions, it rules out strong forms of cross-sectional dependence that might be traced back to the presence of common factors. A variety of factor structures are considered in the panel unit root literature, including cases where either one or both of the common factors and the idiosyncratic components are nonstationary, scenarios of single or multiple factors, and possibilities of serially uncorrelated versus persistent factors (see, e.g., Bai and Ng, 2004; Breitung and Das, 2008; Kapetanios, 2007). These distinct factor structures often deserve specialized defactoring approaches (see,

e.g., Phillips and Sul, 2003; Moon and Perron, 2004; Kapetanios, 2007), which are likely to divert the focus of the present paper on effective centering and standardization in trending and heteroskedastic panels. Accordingly, for the third established feature of macroeconomic panel data, i.e. contemporaneous correlations, we restrict our attention to scenarios of weakly correlated panels. Since $\text{tr}(\Omega_t) = \sum_{i=1}^N \lambda_t^{(i)}$, $\mathcal{A}(ii)$ covers both discrete covariance breaks as well as smoothly trending variances.

2.2 Deterministic terms

Similar to their univariate counterpart, the Dickey-Fuller test (Dickey and Fuller, 1979), PURTs assess the significance of ‘correlation’ among vectors of the first differences ($\Delta \mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$) and the lagged levels of the panel data (\mathbf{y}_{t-1}). Homogeneous PURTs are typically applied to the data after adjustments for higher order serial correlation and deterministic components. Although in practice the prewhitening procedure of removing serial correlation from the data is applied before detrending, we first discuss the detrending issue which is of particular relevance in our case of unit root testing in trending panels. Panel data adjustment for higher order serial correlation is outlined at the end of this section.

Removing the trend in (1) by means of popular schemes such as OLS, GLS or recursive detrending renders homogeneous PURTs to depend on the drift terms in $\boldsymbol{\mu}$, and, hence, requires bias-correction terms. The detrending procedures in Chang (2002) and Breitung and Das (2005) do not require bias adjustment terms as long as the homoskedasticity assumption is maintained. Although both methods affect the pivotalness of PURTs under time-varying volatility, the former allows to trace the adjustment step back to the first and second order moments of a suitably defined test statistic.

The detrending scheme proposed in Chang (2002) involves recursively detrending the lagged level variables to obtain

$$\tilde{\mathbf{y}}_{t-1} = \mathbf{y}_{t-1} + \frac{2}{t-1} \sum_{r=1}^{t-1} \mathbf{y}_r - \frac{6}{t(t-1)} \sum_{r=1}^{t-1} r \mathbf{y}_r. \quad (2)$$

Since $\Delta \mathbf{y}_t$ has non-zero mean, it has to be demeaned. One choice is to center $\Delta \mathbf{y}_t$ in the usual way as

$$\Delta \mathbf{y}_t^* = \Delta \mathbf{y}_t - \frac{1}{T} \sum_{t=2}^T \Delta \mathbf{y}_t, \quad (3)$$

where T in the denominator replaces $T-1$ for notational convenience.

2.3 Heteroskedasticity-robust panel unit root testing for trending panels

2.3.1 Testing with known covariances

For panel unit root diagnosis in heteroskedastic and non-trending panels Herwartz et al. (2016) suggest the statistic¹

$$t_{HSW} = \frac{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \mathbf{y}'_{t-1} \Delta \mathbf{y}_t \Delta \mathbf{y}'_t \mathbf{y}_{t-1}}} \xrightarrow{d} N(0, 1). \quad (4)$$

Evidently, t_{HSW} exhibits a nonzero mean in case of trending panels since, under the null hypothesis, the drift term shows up in $E[\Delta \mathbf{y}_t] \neq 0$. A pivotal test statistic which is robust under the panel unit root, nonzero drift and heteroskedasticity can be obtained after suitable adjustments of both the numerator and the denominator of t_{HSW} . Therefore, before introducing the test statistic in a compact form, first consider a modified version of the numerator of t_{HSW} in (4). With (2) and (3) the summands of the numerator can be rewritten as

$$\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* = \tilde{\mathbf{y}}'_{t-1} \hat{\mathbf{e}}_t = \sum_{q=1}^{t-1} \left(a_{q,t-1} \mathbf{e}'_q \mathbf{e}_t - \frac{1}{T} a_{q,t-1} \sum_{r=2}^T \mathbf{e}'_q \mathbf{e}_r \right), \quad (5)$$

where

$$\Delta \mathbf{y}_t^* = \hat{\mathbf{e}}_t = \mathbf{e}_t - \frac{1}{T} \sum_{t=2}^T \mathbf{e}_t, \quad (6)$$

and finite weighting coefficients $a_{q,t-1}$ read as

$$a_{q,t-1} = 1 + \frac{2}{t-1}(t-q) - 3 \left(1 - \frac{(q-1)q}{(t-1)t} \right). \quad (7)$$

Derived from data detrended according to (2) and (3), the expression in (5) has a non-zero expectation in the absence of homoskedasticity under the null hypothesis of a panel unit root.

To establish asymptotic normality of our test it is convenient to proceed in two steps. Firstly, we outline an infeasible test statistic processing true time varying covariances for which central limit theory is available to establish an asymptotic Gaussian distribution under the panel unit root. Below we refer to this variant as the theoretical version of the test and denote it with τ . Secondly, we consider a feasible variant of the statistic, denoted $\hat{\tau}$ which builds upon estimated second order moments. For this, we prove its asymptotic equivalence with τ , and, hence, the asymptotic Gaussian distribution for the feasible PURT under the null hypothesis.

¹Other heteroskedasticity-robust PURTs have been proposed in Demetrescu and Hanck (2012a,b) and Westerlund (2014). A common limitation of these PURTs, however, is that in the presence of linear trends (i.e., $\boldsymbol{\mu} \neq 0$ in (1)), applying standard detrending schemes does not retain the pivotalness of the tests if the data exhibit variance breaks.

The theoretical version of the new test statistic is a modification of t_{HSW} with adjustments for the non-zero mean in the numerator, and corresponding changes for the variance (in the denominator). Specifically, the test statistic with theoretical moments is given by

$$\tau = \frac{\sum_{t=2}^T \frac{1}{\sqrt{NT}} (\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* - E[\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^*])}{\sqrt{\text{Var} \left[\sum_{t=2}^T \frac{1}{\sqrt{NT}} \tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* \right]}}. \quad (8)$$

While t_{HSW} defined in (4) relies on a White-type covariance estimator, combining the recursive detrending scheme with the mean adjustment invokes dependencies of $\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^*$ in the time dimension. Accordingly, the variance of the numerator in (8) reads as:

$$\begin{aligned} s_{NT}^2 &:= \text{Var} \left[\sum_{t=2}^T \frac{1}{\sqrt{NT}} \tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* \right] \\ &= \frac{1}{NT} \left(E \left[\sum_{t=2}^T \tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* \right]^2 - \left(\sum_{t=2}^T \nu_t \right)^2 \right) = \zeta_1 - \zeta_2 + \zeta_3 + \zeta_4 + \zeta_5 - \frac{1}{NT} \left(\sum_{t=2}^T \nu_t \right)^2. \end{aligned} \quad (9)$$

The expansion of the expectation in (9) yields components ζ_1, \dots, ζ_5 which can be shown to correspond to the following quantities

$$\begin{aligned} \zeta_1 &= \frac{2}{NT} \sum_{q=1}^{T-1} \sum_{r=q+1}^{T-1} \sum_{s=q+1}^T \sum_{t=r+1}^T \bar{a}_{q,s-1} \bar{a}_{r,t-1} (\text{tr}(\Omega_q \Omega_r) + \text{tr}(\Omega_q) \text{tr}(\Omega_r)) \\ \zeta_2 &= \frac{2}{NT} \sum_{q=1}^{T-1} \sum_{s=q+1}^T \sum_{t=q+1}^T \tilde{a}_{q,s-1} \bar{a}_{q,t-1} \text{tr}(\Omega_q \Omega_s) \\ \zeta_3 &= \frac{1}{NT} \sum_{q=1}^{T-1} \sum_{t=q+1}^T \tilde{a}_{q,t-1}^2 \text{tr}(\Omega_q \Omega_t) \\ \zeta_4 &= \frac{1}{NT} \sum_{q=1}^{T-1} \sum_{t=q+1}^T \bar{a}_{q,t-1}^2 \left(E[(\mathbf{e}'_q \mathbf{e}_q)^2] + \sum_{r=1, r \neq q, t}^T \text{tr}(\Omega_q \Omega_r) \right) \\ \zeta_5 &= \frac{2}{NT} \sum_{q=1}^{T-2} \sum_{s=q+1}^{T-1} \sum_{t=s+1}^T \bar{a}_{q,t-1} \bar{a}_{q,s-1} \left(E[(\mathbf{e}'_q \mathbf{e}_q)^2] + \sum_{r=1, r \neq q, t, s}^T \text{tr}(\Omega_q \Omega_r) \right), \end{aligned} \quad (10)$$

where $\tilde{a}_{q,t-1} = (1 - \frac{1}{T}) a_{q,t-1}$ and $\bar{a}_{q,t-1} = \frac{1}{T} a_{q,t-1}$ with coefficients $a_{q,t-1}$ defined in (7).

Similarly,

$$\nu_t := E[\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^*] = - \sum_{q=1}^{t-1} \bar{a}_{q,t-1} \text{tr}(\Omega_q). \quad (11)$$

2.3.2 A feasible robust PURT

The feasible counterpart of τ in (8) is given by

$$\hat{\tau} = \frac{\sum_{t=2}^T \frac{1}{\sqrt{NT}} (\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* - \hat{\nu}_t)}{\hat{s}_{NT}}, \quad (12)$$

where the estimator of ν_t is based on the estimation of the traces of the covariance matrices Ω_q . More precisely, we replace $\text{tr}(\Omega_q)$ by $\hat{e}'_q \hat{e}_q$ where \hat{e}_q is a vector of centered residuals (first differences) as defined in (6). Furthermore, the non-standard form of $\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^*$ renders a straightforward estimation of the variance s_{NT}^2 by $\sum_{t=2}^T \frac{1}{NT} (\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^* - \nu_t)^2$ impossible. We rely on the representation in (9) and replace, similarly to the mean estimator, $\text{tr}(\Omega_q \Omega_r)$ by $\hat{e}'_q \hat{e}_r \hat{e}'_q \hat{e}_r$ and $E[(e'_q e_q)^2]$ by $(\hat{e}'_q \hat{e}_q)^2$. Detailed representations of $\hat{\nu}_t$ and \hat{s}_{NT}^2 are given in the online Supporting Information. The following proposition states the asymptotic normality of the statistic in (12).

Proposition 1. *Under assumptions \mathcal{A} and under the null hypothesis $H_0 : \rho = 1$ the test statistic in (12) is asymptotic normally distributed, i.e., for $N, T \rightarrow \infty$ with $N/T^2 \rightarrow 0$*

$$\hat{\tau} \xrightarrow{d} \mathcal{N}(0, 1). \quad (13)$$

Due to dependence between the terms $\tilde{\mathbf{y}}'_{t-1} \Delta \mathbf{y}_t^*$, central limit theory for serially independent processes is not sufficient to prove asymptotic normality of the test statistic. However, Proposition 1 can be proven based on a central limit theorem for near-epoch dependent sequences. Similar to martingale difference sequences, we allow for a regulated form of serial dependence while conditioning on sets generated by functions of the increments \mathbf{e}_t . The proof is given in the online Supporting Information. As it will turn out, the additional requirement of $N/T^2 \rightarrow 0$ is necessary for $\hat{\tau}$ to fulfill the conditions of the central limit theorem, as well as for applying the test to prewhitened data (Herwartz et al., 2016). The following proposition asserts consistency of the statistic in (12) and states its asymptotic behaviour under the local alternative $H_l : \rho = 1 - c/(T\sqrt{N})$.

Proposition 2. *Under assumptions \mathcal{A} and the local alternative hypothesis $H_l : \rho = 1 - c/(T\sqrt{N})$, the test statistic in (12) is asymptotic normally distributed, i.e., for $N, T \rightarrow \infty$, $N/T^2 \rightarrow 0$*

$$\hat{\tau} \xrightarrow{d} \mathcal{N}(\varphi, 1), \quad (14)$$

with negative expectation $\varphi < 0$.

The proof of Proposition 2 along with an analytical expression for φ is provided in Appendix A.4 of Supporting Information.

2.4 Short-run dynamics

To eliminate short-run serial correlation from the data, prewhitening is an important procedure which leaves the limiting distribution of the tests unaffected (Breitung and Das, 2005). This

procedure requires estimating individual-specific autoregressions of the first differences under H_0 , i.e.,

$$\Delta y_{it} = \sum_{r=1}^{p_i} b_{ir} \Delta y_{i,t-r} + e_{it}. \quad (15)$$

Prewhitened data is then obtained as

$$\hat{y}_{it} = y_{it} - \hat{b}_{i1} y_{i,t-1} - \dots - \hat{b}_{ip_i} y_{i,t-p_i}, \quad (16)$$

and

$$\hat{\Delta} y_{it} = \Delta y_{it} - \hat{b}_{i1} \Delta y_{i,t-1} - \dots - \hat{b}_{ip_i} \Delta y_{i,t-p_i}. \quad (17)$$

Any consistent lag-length selection criterion can be applied to decide upon the lag orders p_i . In cases where both short-run dynamics and deterministic patterns are present in the data, prewhitening should precede detrending. The prewhitening regression should include an intercept term if the model features linear time trends under the alternative hypothesis. While the choice of lag selection criteria may not necessarily affect asymptotic properties of unit root tests, it might however impact on finite sample performances of tests in the presence of heteroskedasticity. Cavaliere et al. (2015) have shown that standard lag selection criteria tend to over-fit the lag order in the presence of heteroskedasticity and, hence, induce power losses in the (wild bootstrap version of the) univariate augmented Dickey-Fuller tests. Their modified lag selection criteria, which rescale the data by an estimate of the underlying time-varying volatility process, significantly mitigate the power losses while retaining comparative size properties. However, it is noteworthy that estimating time-varying volatility is more challenging in the panel data framework as each time time series could have distinct underlying volatility processes. Moreover, the effect of potential power losses on unit root testing are likely less severe in PURTs than in univariate tests as the former additionally exploit the cross-sectional dimension of the data. Hence, we focus on the asymptotic behavior of the proposed test, and do not address potential power losses to our test that might be induced by the use of (slightly) over-fitting models employed for prewhitening.

3 Monte Carlo study

3.1 The simulation design

To evaluate the finite sample properties of the proposed test $\hat{\tau}$, we consider the following DGPs taken from Pesaran (2007):

$$\text{DGP1: } \mathbf{y}_t = \boldsymbol{\mu} + (\mathbf{j} - \boldsymbol{\rho}) \odot \boldsymbol{\beta} t + \boldsymbol{\rho} \odot \mathbf{y}_{t-1} + \mathbf{e}_t, \quad t = -50, \dots, T, \quad (18)$$

$$\text{DGP2: } \mathbf{y}_t = \boldsymbol{\mu} + (\mathbf{j} - \boldsymbol{\rho}) \odot \boldsymbol{\beta} t + \boldsymbol{\rho} \odot \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t = \mathbf{b} \odot \boldsymbol{\epsilon}_{t-1} + \mathbf{e}_t, \quad (19)$$

where bold entries indicate vectors of dimension $N \times 1$, \mathbf{j} is a vector of ones and \odot denotes the Hadamard product. The DGP1 formalizes AR(1) models with serially uncorrelated innovations while DGP2 introduces AR(1) disturbances. Both DGPs formalize a panel random walk with drift under the null hypothesis, and a panel of trend stationary processes with individual effects under the alternative. Empirical size obtains by setting $\boldsymbol{\rho} = \mathbf{j}$ and power is simulated as $\boldsymbol{\rho} = 0.9\mathbf{j}$. Results for DGPs with heterogeneous autoregressive coefficients under the alternative hypothesis, i.e., $\boldsymbol{\rho} = (\rho_1, \dots, \rho_N)$, $\rho_i \sim \text{iid } U(0.85, 0.95)$, are qualitatively identical.² Individual effects, trend parameters as well as serial correlation of innovations are modeled as in Pesaran (2007): $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)'$, $\mu_i \sim \text{iid } U(0, 0.02)$ and $\mathbf{b} = (b_1, \dots, b_N)'$, $b_i \sim \text{iid } U(0.2, 0.4)$.

To separate the issue of cross-sectional correlation from variance breaks, we employ the decomposition

$$\Omega_t = \Phi_t^{1/2} \Psi \Phi_t^{1/2},$$

where $\Phi_t = \text{diag}(\sigma_{1t}^2, \dots, \sigma_{Nt}^2)$ and Ψ is a (time invariant) correlation matrix characterizing Ω_t . Cross-sectional independence is obtained by setting Ψ to an identity matrix of order N . We generate a weak form of cross-sectional correlation by means of the spatial autoregressive (SAR) error structure used in Herwartz and Siedenburg (2008). Specifically, we take Ψ_{SAR} that is implied by the SAR model

$$\mathbf{e}_t = (I_N - \Theta W)^{-1} \boldsymbol{\xi}_t, \quad \text{with } \Theta = 0.8 \quad \text{and} \quad \boldsymbol{\xi}_t \sim \text{iid } N(\mathbf{0}, I_N),$$

where W is the so-called spatial weights matrix. In this particular case, W is a row normalized symmetric contiguity matrix of the ‘*g ahead and g behind*’ structure, with $g = 1$ (see, e.g., Kelejian and Prucha, 1999). The resulting covariance matrix of \mathbf{e}_t is given by $\Omega_{SAR} = ((I_N - \Theta W)'(I_N - \Theta W))^{-1}$, and Ψ_{SAR} is the correlation matrix implied by Ω_{SAR} .

Cross-section specific volatility shifts are generated as

$$\sigma_{it}^2 = \begin{cases} \sigma_{i1}^2, & \text{if } t < \lfloor \gamma_i T \rfloor, \quad (0 < \gamma_i < 1) \\ \sigma_{i2}^2, & \text{otherwise,} \end{cases}$$

where γ_i refers to the time a variance break occurs and $\lfloor \gamma_i T \rfloor$ denotes the integer part of $\gamma_i T$. In the homoskedastic case, $\sigma_{i1} = \sigma_{i2} = 1$. We introduce heteroskedasticity by changing the post-break variance to $\sigma_{i2} = 1/3$, for a negative variance break, and to $\sigma_{i2} = 3$, for a positive one. Regarding the timing of the variance breaks, we consider scenarios of homogeneously early ($\gamma_i = 0.2$) or late ($\gamma_i = 0.8$) variance breaks for all panel units. Main findings of the

²Moreover, recent papers, e.g., Homm and Breitung (2012), also consider power against explosive alternatives ($\rho > 1$). Using a right-sided testing, the proposed test $\hat{\tau}$ is powerful against the alternative that $\boldsymbol{\rho} = 1.03\mathbf{j}$, even for $T = 25$. The corresponding simulation results are available upon request.

simulation exercise remain qualitatively unaffected by consideration of randomly distinct break points $\gamma_i \sim \text{iid} U(0.1, 0.9)$. Data are generated for all combinations of $N \in [50, 100, 250]$ and $T \in [25, 50, 100, 250]$. To mitigate the potential impacts of initial values on our analysis, we generate and discard 50 presample observations.

3.2 Simulation results

In the following we discuss simulation results on the finite sample performance of the proposed test statistic $\hat{\tau}$. Simulation results for t_{HSW} in (4) and the test suggested by Demetrescu and Hanck (2012a) are provided to highlight the risks of using existing heteroskedasticity-robust tests for trending panels. The latter test reads as

$$t_{DH} = \frac{\sum_{t=1}^T \text{sgn}(\mathbf{y}_{t-1})' \Delta \mathbf{y}_t}{\sqrt{\sum_{t=1}^T \text{sgn}(\mathbf{y}_{t-1})' \Delta \mathbf{y}_t \Delta \mathbf{y}_t' \text{sgn}(\mathbf{y}_{t-1})}} \xrightarrow{d} N(0, 1), \quad (20)$$

where $\text{sgn}(\cdot)$ denotes the sign function. Similar to t_{HSW} , t_{DH} should be applied on detrended data.

For the new test $\hat{\tau}$, we also document results for its theoretical counterpart τ determined from the true covariance matrices Ω_t (see (8)). Presenting simulation results for both $\hat{\tau}$ and τ is meant to highlight finite sample performance of $\hat{\tau}$ that can be traced back to the use of moment estimators.

3.2.1 Cross-sectionally independent panels

Simulation results for data generated according to DGP1 and cross-sectionally independent panels are documented in Table 1. Results in the upper panel of this table show that, under homoskedasticity, the recursive detrending scheme in Chang (2002) leaves the pivotalness of heteroskedasticity-robust tests unaffected. With respect to rejection frequencies under the alternative hypothesis, it can be seen that using estimated covariance matrices induces considerable power loss under a small time dimension $T = 25$. However, this power loss vanishes with increasing T . Furthermore, the new test $\hat{\tau}$ is generally as powerful as t_{HSW} and more powerful than t_{DH} . Hence, it is worthwhile noting that our adjustment for obtaining robustness to time-varying volatility does not come at a cost of reduced power. In view of the fact that the reported empirical powers are not size adjusted, the power estimates for $\hat{\tau}$ are rather remarkable.

When early negative variance breaks are introduced, t_{HSW} and t_{DH} display zero rejection frequencies under the null hypothesis. On the contrary, $\hat{\tau}$ holds remarkable size control, except for small T ($T = 25$) where it is substantially undersized. These size distortions, however,

improve markedly as the time dimension increases to $T = 50$. The new test also has significant power under early variance breaks although it is less than the power under homoskedasticity. In comparison with $\hat{\tau}$, the White-type tests t_{HSW} and t_{DH} have substantially weaker power, with both tests showing almost zero probability of rejecting the alternative hypothesis until the time dimension increases to $T = 100$.

Size distortions of t_{HSW} and t_{DH} are also observed when a late positive volatility shift is considered, but this time with huge oversizings. On the contrary, $\hat{\tau}$ displays a fairly good size precision. Consistent with results in Herwartz et al. (2016) for non-trending data, power seems to be unaffected by late positive variance breaks but reduced by early negative volatility shifts. In general, simulation results documented in Table 1 demonstrate not only the risk of using t_{HSW} and t_{DH} for trending time series, but also the satisfactory finite sample performance of $\hat{\tau}$ for trending heteroskedastic data.

Place Table 1 here.

3.2.2 DGPs with cross-sectionally correlated panels

The left-hand side block of Table 2 documents simulation results for $\hat{\tau}$ applied on data generated according to DGP1 for weakly correlated panels. Results available upon request show that size distortions of t_{HSW} and t_{DH} observed for cross-sectionally independent panels (Table 1) carry over to panels with weak forms of cross-sectional correlation. Hence, we focus on the implications of cross-sectional correlation for the new test $\hat{\tau}$. Confirming the asymptotic considerations, a relatively larger cross-sectional dimension N is required for the empirical size of $\hat{\tau}$ to come closer to the nominal significance levels. Moreover, the statistic $\hat{\tau}$ is less powerful under the SAR(1) model than under independent panels—a result consistent with those documented in Herwartz et al. (2016) for non-trending series.

3.2.3 DGPs with serially correlated innovations

To evaluate how the proposed test $\hat{\tau}$ performs for data with serially correlated disturbances, we generate data according to DGP2 in (19) and subject it to prewhitening before detrending. The corresponding simulation results are documented in the middle- and right-hand side blocks of Table 2. The results show that serial correlation and the ensuing prewhitening procedure entail marked size distortions for small time dimensions. This result could be explained by

noting that estimation errors arising from the prewhitening procedure introduce finite sample correlations between the lagged level and first differenced series, thereby inducing a non-zero mean to the numerator of the test statistic in (12). However, size distortions vanish as T grows, and empirical power grows in T and N .

Place Table 2 here.

Place Table 3 here.

3.2.4 Testing against a heterogeneous alternative and local power

As documented in Table 2 scenarios of early negative variance shifts are most challenging for the detection of stationary panels. For this particular scenario of weakest simulated power, Table 3 documents rejection frequencies for $\hat{\tau}$ if only fractions of 20%, 50% and 90% of the panel members accord with panel stationarity, while the remaining panel members are characterized by stochastically trending data. Not surprisingly, under the heterogeneous alternative the test suffers from some power loss in comparison with simulation results documented in Table 2. Pointing to consistency of $\hat{\tau}$, however, the documented rejection frequencies increase throughout with the sample sizes. Moreover, power estimates are more favourable, the larger is the fraction of panel members that accord with the alternative hypothesis.

Completing our Monte-Carlo analysis, Figure 1 shows local power figures (mostly) for $\hat{\tau}$ under alternative covariance scenarios. Local power estimates for t_{HSW} under homoskedasticity are also provided for purposes of comparison. Irrespective of the second order (break) properties of model residuals, local power estimates point to the consistency of the new test. As a confirmation of the tabulated simulation results, scenarios with late positive variance shift obtain a more favourable performance pattern under the local alternative in comparison with early negative variance breaks. Interestingly, in case of homoskedastic trending panels t_{HSW} and $\hat{\tau}$ exhibit very similar local power properties, although the latter statistic is computationally much more demanding.

Place Figure 1 here.

3.2.5 Summary of simulation results

The simulation results reported in Table 1 show that existing heteroskedasticity-robust PURTs exhibit huge size distortions (either undersizing or oversizing) when applied to detrended data with time-varying volatility. The proposed test, however, performs remarkably well in this scenario. Results documented in Table 2 show that the new test has fairly good finite sample properties even when the data are not only trending and heteroskedastic, but also cross-sectionally and serially correlated. Therefore, the new test should be helpful in (often complex) empirical applications. Rejections of the null hypothesis of the panel unit root could indicate that only a significant fraction of the cross section is characterized by a stationary process. Asymptotically the suggested test statistic exhibits non-trivial power features in close neighbourhoods of the null hypothesis. However, results not reported here for space considerations show that $\hat{\tau}$ does not remain pivotal under strong forms of cross-sectional dependence such as factor structures (Pesaran, 2007). An effective way of panel unit root testing under strong forms of cross-sectional correlation is to remove the common factor from the data (see for example Bai and Ng, 2004 and Moon and Perron, 2004). While the test in Westerlund (2014) uses this approach, it is, however, not pivotal in the presence of linear trends.

4 Is energy use per capita trend or difference stationary?

4.1 Background

Whether energy use per capita is trend or difference stationary has been intensively investigated in the past two decades. The growing interest in testing the stationarity of per capita energy consumption is attributed to three main reasons (e.g., Hsu et al., 2008; Narayan and Smyth, 2007). First, knowing the direction of causality between per capita energy use and economic growth has gained significant policy relevance as it has direct implications on governments' involvement in global efforts to reduce greenhouse gas emissions. On the one hand, if causality runs from energy consumption to growth, reductions in energy use will have adverse effects on economic growth and, hence, generates reluctance on the part of policy makers to commit to substantial energy use reductions. On the other hand, if causality runs from growth to energy use, and not vice versa, reductions in energy consumption will not be harmful for economic growth. The order of integration of energy use per capita has implications on testing and interpreting the relationship between energy use and GDP per capita. For instance, Granger predictability tests employing level vector autoregressions could be misleading if the series

are nonstationary and not cointegrated. Conversely, Granger predictability testing by means of variables in levels will be appropriate if the series are either stationary or cointegrated. Consequently, unit root testing is routinely performed before testing for cointegration between energy use and GDP per capita.

Second, stationarity of energy use per capita has implications for the effectiveness of energy policies such as import tariffs on fuels and vehicles or carbon taxes on transportation fuels. In particular, if energy consumption is a stationary process, it will return to its trend after a policy shock. This implies that energy saving policies will have transitory effects only. On the other hand, if energy consumption contains a unit root, such policies will have a permanent impact. Furthermore, nonstationarity implies that (permanent) shocks to energy use are more likely to affect other sectors of the economy as well as macroeconomic aggregates (Narayan and Smyth, 2007).

Third, the order of integration of energy consumption has implications for forecasting energy demand. For instance, if energy consumption is trend stationary, its past behaviour offers valuable information to forecast future energy demand. However, if energy consumption is a unit root process, it does not follow a predictable path and, hence, forecasting energy demand will be more difficult than in the stationary case.

Efforts to test for a unit root in energy use per capita have initially relied on univariate tests. Hsu et al. (2008) review the empirical literature on unit root testing of energy use per capita. Most studies, including Glasure and Lee (1998), Beenstock et al. (1999) and McAviney and Yannopoulos (2003) report that the null hypothesis of an $I(1)$ energy consumption series can not be rejected at conventional levels of significance. As an exception to this general conclusion, Altinay and Karagol (2004) document evidence in favor of characterizing energy use in Turkey during 1950–2000 as a trend stationary process. However, given the low power of univariate tests in finite samples, it is not clear if the failure to reject the null of a unit root is an evidence of a truly $I(1)$ series. To circumvent this problem, a few studies have recently applied PURTs to examine the stationarity of energy use per capita. Results have been generally mixed, however. For instance, Joyeux and Ripple (2007) employ the PURTs suggested in Levin et al. (2002) and Im et al. (2003) and find that energy consumption measures are $I(1)$. Narayan and Smyth (2007), on the other hand, report that the unit root null hypothesis can be rejected at the 10% level of significance for 56 of the 182 countries they considered. However, they find strong evidence of a (trend) stationary energy consumption by employing the PURT of Im et al. (2003). Nevertheless, these results should be seen with caution as the studies employ standard PURTs, which are not pivotal if the series exhibit volatility shifts.

4.2 Panel unit root test results

In this section, we study the order of integration of energy consumption per capita using the test suggested in this paper, $\hat{\tau}$, vis-a-vis the heteroskedasticity-robust tests of Herwartz et al. (2016) and Demetrescu and Hanck (2012a) lacking applicability for such trending data. We analyse annual data of energy use per capita (kilogram of oil equivalent per capita) obtained from World Development Indicators (World Bank, 2016). In this data set, energy use refers to “*use of primary energy before transformation to other end-use fuels, which is equal to indigenous production plus imports and stock changes, minus exports and fuels supplied to ships and aircraft engaged in international transport.*” The study covers 23 OECD economies (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom and the United States) from 1960 to 2014. The economies are selected according to data availability. As transforming the series into natural logarithms before undertaking unit root testing is a standard practice in the literature, we test for unit roots both on original series as well as their logarithmic values.

To get an impression if variances in the energy use per capita series exhibit significant changes over time, we plot variance profiles in Figure 2. Variance profiles $\hat{v}_i(w)$ are computed as

$$\hat{v}_i(w) = \frac{\sum_{t=1}^{\lfloor sT \rfloor} \hat{\eta}_{it}^2 + (wT - \lfloor wT \rfloor) \hat{\eta}_{i[\lfloor wT \rfloor + 1]}^2}{\sum_{t=1}^T \hat{\eta}_{it}^2}, 0 \leq w \leq 1, \quad (21)$$

where the $\hat{\eta}_{it}$'s are obtained as residuals from AR(1) regressions of the series. Plotting $\hat{v}_i(w)$ against w , it is straightforward to see that a homoskedastic series would fall on the 45° line and deviations from the diagonal indicate time-varying variances. Figure 2 reveals that time-varying variances characterize energy per capita series in most cross section members.

Place Figure 2 here.

Place Table 4 here.

Panel unit root test results are reported in Table 4. Results for all the tests overwhelmingly show that energy use per capita has a unit root. This evidence is consistent with the findings of most of the empirical studies on the area, except, e.g., Narayan and Smyth (2007). However,

it is well-known that unit root test results often depend on the specific time period under study. To address this caveat, we perform panel unit root testing on rolling windows of 40 years. Corresponding results depicted in Figure 3 show that while energy use per capita is difference stationary for most of the period, it could be considered trend stationary—at least at the 10 percent significance level—for the sample periods starting between 1965 and 1968. It is worthwhile noting that $\hat{\tau}$ has the lowest p -value of the three tests in almost all the considered periods and could suggest an inferential outcome which is distinct from that of the other two tests. In particular, for the period spanning 1966-2005 and based on the 5 percent significance level, $\hat{\tau}$ implies that log energy per capita series can be considered trend stationary while the other two tests suggest to treat the series as difference stationary. Moreover, our results also highlight the risk of deciding on stationarity of series using one specific time window.

Place Figure 3 here.

5 Conclusions

In this paper, we suggested a new panel unit root test (PURT) for the empirically-relevant case where trending panels exhibit time-varying volatility. The test makes use of the recursive detrending scheme proposed in Chang (2002), and the construction of the test statistic fully accounts for non-zero expectation of the pooled panel regression estimator and the variance of its centred counterpart. Accordingly, the resulting test statistic has a Gaussian limiting distribution. Monte Carlo simulation results show that the test has satisfactory finite sample properties. In particular, the test tends to be conservative, while it shows remarkable power. Hence, this test should be useful in panel unit root testing of several trending macroeconomic and financial time series such as GDP per capita, industrial production, money supply and commodity prices.

The empirical illustration examined the order of integration of energy use per capita. Results using data from 23 OECD economies for the period 1960-2014 show that energy use per capita is often difference stationary. Yet, there are also a few sub-periods for which the series could be considered as trend stationary.

A particular limitation of the suggested test is that it does not perform well under a strong form of cross-sectional dependence. An effective way of panel unit root testing under strong forms of cross-sectional correlation is to remove the common factor from the data (Bai and Ng, 2004; Moon and Perron, 2004; Kapetanios, 2007). Consequently, it appears worthwhile to see in a future research if such an approach would yield a panel unit root test that works for strongly correlated panels with trending and heteroskedastic time series.

Data Availability Statement

The data used in this study are openly available from World Development Indicators of the World Bank: <https://data.worldbank.org>. Accessed on September 23, 2016.

References

- Altınay, G. and Karagol, E. (2004). Structural break, unit root, and the causality between energy consumption and GDP in Turkey. *Energy Economics*, 26(6):985–994.
- Anselin, L. (2013). *Spatial Econometrics: Methods and Models*, volume 4. Springer Science & Business Media.
- Bai, J. and Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, 72(4):1127–1178.
- Beenstock, M., Goldin, E., and Nabot, D. (1999). The demand for electricity in Israel. *Energy Economics*, 21(2):168–183.
- 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*, 24:88–108.
- 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.
- Cavaliere, G. (2004). Unit root tests under time varying variances. *Econometric Reviews*, 23(4):259–292.
- Cavaliere, G., Phillips, P. C. B., Smeekes, S., and Taylor, A. M. R. (2015). Lag length selection for unit root tests in the presence of nonstationary volatility. *Econometric Reviews*, 34(4):512–536.
- Cavaliere, G. and Taylor, A. M. R. (2007). Testing for unit roots in time series with non-stationary volatility. *Journal of Econometrics*, 140(2):919–947.
- Chang, Y. (2002). Nonlinear IV unit root test in panels with cross sectional dependency. *Journal of Econometrics*, 110(2):261–292.
- Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians*. Oxford University Press.
- Demetrescu, M. and Hanck, C. (2012a). A simple nonstationary-volatility robust panel unit root test. *Economics Letters*, 117(2):10–13.

- Demetrescu, M. and Hanck, C. (2012b). Unit root testing in heteroscedastic panels using the Cauchy estimator. *Journal of Business and Economic Statistics*, 30(2):256–264.
- 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*, 74:427–431.
- Glasure, Y. U. and Lee, A.-R. (1998). Cointegration, error-correction, and the relationship between GDP and energy: The case of South Korea and Singapore. *Resource and Energy Economics*, 20(1):17–25.
- Hamori, S. and Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift. *Economics Letters*, 57(3):245–253.
- Herwartz, H., Maxand, S., Raters, F. H., Walle, Y. M., et al. (2018). Panel unit-root tests for heteroskedastic panels. *Stata Journal*, 18(1):184–196.
- Herwartz, H. and Siedenburg, F. (2008). Homogenous panel unit root tests under cross sectional dependence: Finite sample modifications and the wild bootstrap. *Computational Statistics and Data Analysis*, 53(1):137–150.
- Herwartz, H., Siedenburg, F., and Walle, Y. M. (2016). Heteroskedasticity robust panel unit root testing under variance breaks in pooled regressions. *Econometric Reviews*, 35(5):727–750.
- Homm, U. and Breitung, J. (2012). Testing for speculative bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Econometrics*, 10(1):198–231.
- Hsu, Y.-C., Lee, C.-C., and Lee, C.-C. (2008). Revisited: Are shocks to energy consumption permanent or temporary? New evidence from a panel SURADF approach. *Energy Economics*, 30(5):2314–2330.
- Im, K. S., Pesaran, H. M., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.
- Joyeux, R. and Ripple, R. D. (2007). Household energy consumption versus income and relative standard of living: A panel approach. *Energy Policy*, 35(1):50–60.
- Kapetanios, G. (2007). Dynamic factor extraction of cross-sectional dependence in panel unit root tests. *Journal of Applied Econometrics*, 22(2):313–338.
- Kelejian, H. H. and Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2):509–533.

- 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.
- McAvinchey, I. D. and Yannopoulos, A. (2003). Stationarity, structural change and specification in a demand system: The case of energy. *Energy Economics*, 25(1):65–92.
- Moon, H. R. and Perron, B. (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics*, 122(1):81–126.
- Narayan, P. K. and Smyth, R. (2007). Are shocks to energy consumption permanent or temporary? Evidence from 182 countries. *Energy Policy*, 35(1):333–341.
- 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. (2012). On the interpretation of panel unit root tests. *Economics Letters*, 116(3):545–546.
- Phillips, P. C. B. and Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, 6(1):217–259.
- Sensier, M. and Dijk, D. v. (2004). Testing for volatility changes in US macroeconomic time series. *Review of Economics and Statistics*, 86(3):833–839.
- Westerlund, J. (2014). Heteroscedasticity robust panel unit root tests. *Journal of Business & Economic Statistics*, 32(1):112–135.
- Westerlund, J. (2015). The effect of recursive detrending on panel unit root tests. *Journal of Econometrics*, 185(2):453–467.
- Westerlund, J. and Breitung, J. (2013). Lessons from a decade of IPS and LLC. *Econometric Reviews*, 32(5-6):547–591.
- World Bank (2016). World Development Indicators. <https://data.worldbank.org>. Accessed: September 23, 2016.

Table 1: Empirical rejection frequencies, cross-sectionally independent panels

N	T	5%								10%							
		size				power				size				power			
		τ	$\hat{\tau}$	HSW	DH	τ	$\hat{\tau}$	HSW	DH	τ	$\hat{\tau}$	HSW	DH	τ	$\hat{\tau}$	HSW	DH
<i>Constant variance (HOM)</i>																	
50	25	5.6	4.6	4.8	4.5	40.5	32.0	31.1	19.4	11.1	10.0	9.8	9.9	55.2	49.9	48.3	32.3
50	50	4.6	5.5	4.8	4.3	98.9	98.6	97.4	77.2	10.4	11.1	9.7	9.0	99.6	99.7	99.4	87.7
50	100	4.6	4.3	3.5	4.0	100.0	100.0	100.0	100.0	9.7	9.7	8.4	8.6	100.0	100.0	100.0	100.0
50	250	3.8	4.2	3.6	4.0	100.0	100.0	100.0	100.0	9.0	9.9	8.6	8.6	100.0	100.0	100.0	100.0
100	25	5.4	3.4	4.3	4.7	63.8	50.8	49.6	28.9	11.4	8.2	9.5	9.5	76.1	68.4	67.8	45.6
100	50	5.2	5.5	4.7	4.7	100.0	100.0	100.0	96.2	11.0	10.8	9.7	9.9	100.0	100.0	100.0	98.8
100	100	4.5	4.9	3.7	4.4	100.0	100.0	100.0	100.0	9.4	10.2	8.8	8.7	100.0	100.0	100.0	100.0
100	250	4.9	4.6	3.7	4.4	100.0	100.0	100.0	100.0	10.3	10.0	8.4	9.0	100.0	100.0	100.0	100.0
250	25	5.4	2.0	4.5	4.8	92.7	81.0	83.2	54.2	10.8	5.9	9.5	9.7	96.1	92.1	93.4	70.1
250	50	4.9	4.1	3.8	4.2	100.0	100.0	100.0	100.0	10.4	9.5	9.0	9.2	100.0	100.0	100.0	100.0
250	100	4.7	4.7	3.6	4.1	100.0	100.0	100.0	100.0	9.6	9.5	8.0	8.7	100.0	100.0	100.0	100.0
250	250	4.6	5.0	3.9	4.0	100.0	100.0	100.0	100.0	10.1	10.3	8.4	9.0	100.0	100.0	100.0	100.0
<i>Early negative variance shift (NEG)</i>																	
50	25	4.9	1.4	0.0	0.0	9.6	4.3	0.0	0.0	10.2	4.4	0.0	0.0	17.3	12.1	0.0	0.0
50	50	5.1	3.0	0.0	0.0	48.8	48.2	0.0	0.1	10.7	7.9	0.0	0.0	63.8	66.2	0.1	0.6
50	100	4.7	4.0	0.0	0.0	99.9	99.9	37.8	36.1	10.0	8.9	0.0	0.0	100.0	100.0	54.4	51.4
50	250	4.6	4.5	0.0	0.0	100.0	100.0	100.0	100.0	9.9	10.2	0.0	0.0	100.0	100.0	100.0	100.0
100	25	5.4	0.4	0.0	0.0	12.5	3.3	0.0	0.0	10.5	2.6	0.0	0.0	22.3	12.3	0.0	0.0
100	50	5.2	2.3	0.0	0.0	75.5	72.4	0.0	0.0	10.6	6.0	0.0	0.0	85.8	87.1	0.0	0.1
100	100	4.9	3.9	0.0	0.0	100.0	100.0	57.0	57.5	10.6	8.9	0.0	0.0	100.0	100.0	73.7	73.0
100	250	5.2	4.5	0.0	0.0	100.0	100.0	100.0	100.0	10.5	10.3	0.0	0.0	100.0	100.0	100.0	100.0
250	25	5.0	0.0	0.0	0.0	18.6	0.8	0.0	0.0	10.1	0.4	0.0	0.0	29.2	6.9	0.0	0.0
250	50	5.0	1.0	0.0	0.0	97.8	96.1	0.0	0.0	10.7	4.4	0.0	0.0	99.2	98.9	0.0	0.0
250	100	5.0	3.4	0.0	0.0	100.0	100.0	84.4	87.1	10.1	7.9	0.0	0.0	100.0	100.0	93.4	94.0
250	250	5.0	4.8	0.0	0.0	100.0	100.0	100.0	100.0	10.0	9.3	0.0	0.0	100.0	100.0	100.0	100.0
<i>Late positive variance shift (POS)</i>																	
50	25	4.6	1.6	17.7	11.0	57.7	35.6	77.8	50.8	9.9	7.0	30.7	21.3	73.6	64.3	92.3	68.4
50	50	3.3	3.2	26.0	13.1	99.0	98.8	99.8	95.4	8.2	8.2	39.6	23.4	99.9	99.9	100.0	98.7
50	100	3.1	3.3	21.0	12.2	100.0	100.0	100.0	100.0	7.6	8.4	35.0	21.9	100.0	100.0	100.0	100.0
50	250	2.6	2.7	18.4	12.3	100.0	100.0	100.0	100.0	7.3	8.0	32.1	22.4	100.0	100.0	100.0	100.0
100	25	4.7	0.5	29.1	15.6	85.5	55.6	92.7	72.9	10.0	3.7	44.7	27.0	92.6	83.4	99.0	86.9
100	50	4.6	2.9	34.4	17.4	100.0	100.0	100.0	99.9	9.4	7.8	50.8	29.6	100.0	100.0	100.0	100.0
100	100	3.5	2.9	30.0	16.3	100.0	100.0	100.0	100.0	8.0	7.2	46.7	28.1	100.0	100.0	100.0	100.0
100	250	3.4	3.1	28.8	15.4	100.0	100.0	100.0	100.0	8.3	8.1	45.3	26.8	100.0	100.0	100.0	100.0
250	25	4.4	0.1	57.3	29.4	99.7	85.0	99.7	94.6	10.4	1.1	72.3	44.1	99.9	98.3	100.0	98.8
250	50	3.4	1.9	56.0	27.7	100.0	100.0	100.0	100.0	9.0	6.3	73.7	42.0	100.0	100.0	100.0	100.0
250	100	4.0	3.7	53.7	26.8	100.0	100.0	100.0	100.0	9.0	8.7	71.8	41.7	100.0	100.0	100.0	100.0
250	250	3.1	3.3	57.3	29.1	100.0	100.0	100.0	100.0	7.4	8.1	72.3	44.0	100.0	100.0	100.0	100.0

Notes: τ , $\hat{\tau}$, HSW and DH refer to the PURT statistics given in (8), (12), (4) and (20) respectively. Power is not size adjusted. All results are based on 5000 replications. Data is generated according to DGP1 in (18) and all tests are computed on detrended data.

Table 2: Empirical rejection frequencies of $\hat{\tau}$, diverse scenarios

N	T	DGP1, SAR(1) model				DGP2, Independence				DGP2, SAR(1) model			
		5%		10%		5%		10%		5%		10%	
		size	power	size	power	size	power	size	power	size	power	size	power
<i>Constant variance (HOM)</i>													
50	25	4.3	16.2	11.2	28.9	0.0	0.0	0.0	0.1	0.0	0.1	0.5	0.7
50	50	3.8	60.4	10.3	78.0	0.1	29.7	0.8	50.3	0.2	8.1	1.2	20.6
50	100	3.6	100.0	9.3	100.0	1.3	100.0	3.7	100.0	1.2	95.4	4.1	98.7
50	250	3.2	100.0	8.6	100.0	2.9	100.0	6.7	100.0	1.9	100.0	5.5	100.0
100	25	4.8	24.7	10.6	40.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2
100	50	4.4	89.7	10.5	95.6	0.1	58.7	0.3	77.6	0.1	17.7	0.9	36.8
100	100	4.2	100.0	10.0	100.0	1.0	100.0	3.0	100.0	1.0	100.0	4.0	100.0
100	250	3.8	100.0	9.2	100.0	2.5	100.0	6.0	100.0	2.2	100.0	6.4	100.0
250	25	3.7	42.6	8.9	61.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
250	50	4.5	99.9	9.9	100.0	0.0	94.0	0.1	98.3	0.1	48.1	0.6	68.4
250	100	4.1	100.0	9.1	100.0	0.4	100.0	1.3	100.0	0.8	100.0	2.9	100.0
250	250	4.5	100.0	10.2	100.0	2.6	100.0	5.9	100.0	2.5	100.0	6.4	100.0
<i>Early negative variance shift (NEG)</i>													
50	25	2.9	5.1	7.7	13.7	0.0	0.0	0.0	0.0	0.1	0.0	0.5	0.3
50	50	3.3	20.8	8.9	38.1	0.1	0.4	0.4	2.0	0.3	0.9	1.4	3.0
50	100	3.7	79.2	9.6	90.4	0.7	87.7	3.0	95.9	0.8	30.3	3.1	55.3
50	250	3.5	100	8.5	100.0	3.0	100.0	8.0	100.0	1.7	100.0	6.3	100.0
100	25	2.0	4.8	5.7	13.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
100	50	3.3	35.3	8.5	53.4	0.0	0.2	0.1	0.7	0.2	0.5	0.6	2.1
100	100	4.6	97.7	10.1	99.3	0.6	99.6	2.0	100.0	0.9	66.3	2.9	84.8
100	250	4.4	100.0	9.5	100.0	3.7	100.0	8.5	100.0	2.3	100.0	6.0	100.0
250	25	0.6	3.6	3.1	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
250	50	3.0	61.9	6.9	78.1	0.0	0.0	0.0	0.1	0.0	0.2	0.0	0.9
250	100	4.2	100.0	9.2	100.0	0.2	100.0	1.2	100.0	0.5	97.9	2.0	99.5
250	250	4.6	100.0	9.8	100.0	3.9	100.0	9.5	100.0	2.5	100.0	7.6	100.0
<i>Late positive variance shift (POS)</i>													
50	25	1.8	14.8	7.9	34.7	0.0	0.3	0.4	4.4	0.2	0.7	1.5	4.5
50	50	1.8	56.7	7.1	80.1	0.8	63.1	4.7	83.6	0.6	16.4	3.8	37.9
50	100	1.4	99.0	6.0	99.8	2.2	100.0	6.6	100.0	0.9	85.8	4.8	96.8
50	250	1.1	100.0	5.6	100.0	2.8	100.0	7.1	100.0	0.9	100.0	4.9	100.0
100	25	2.0	24.8	7.2	50.0	0.0	0.2	0.2	3.2	0.1	0.5	0.8	4.4
100	50	2.6	89.8	7.6	97.0	1.2	91.9	4.4	98.4	1.0	39.9	4.1	66.0
100	100	2.6	100.0	7.7	100.0	3.0	100.0	8.2	100.0	1.6	99.7	5.8	100.0
100	250	2.2	100.0	7.0	100.0	3.1	100.0	8.3	100.0	2.0	100.0	6.0	100.0
250	25	0.8	48.1	4.3	77.4	0.0	0.0	0.0	1.1	0.0	0.2	0.1	2.9
250	50	3.1	99.9	8.0	100.0	0.4	100.0	3.7	100.0	0.8	85.0	4.2	95.7
250	100	3.0	100.0	8.2	100.0	3.1	100.0	9.0	100.0	2.1	100.0	6.3	100.0
250	250	2.8	100.0	7.7	100.0	4.1	100.0	10.1	100.0	2.5	100.0	7.8	100.0

Notes: Data is generated according to DGP1 in (18) for results in the left-hand side block, while DGP2 in (19) is used to generate data for results documented in the middle and right-hand side blocks of the table. Testing is performed on detrended data. For DGP2, detrending is preceded by prewhitening. Power is not size adjusted and all results are based on 5000 replications.

Table 3: Power of $\hat{\tau}$ when a subset of panel units are stationary

N	T	DGP1, SAR(1) model				DGP2, Independence				DGP2, SAR(1) model			
		α				α				α			
		0	0.2	0.5	0.9	0	0.2	0.5	0.9	0	0.2	0.5	0.9
<i>Early negative variance shift (NEG)</i>													
50	25	2.9	3.0	3.6	4.7	0.0	0.0	0.0	0.0	0.3	0.1	0.2	0.2
50	50	3.3	4.7	8.5	18.0	0.3	0.4	0.7	1.6	0.5	0.5	0.7	1.5
50	100	3.7	8.8	24.7	66.3	1.3	5.4	27.3	84.8	1.0	2.1	7.0	29.6
50	250	2.8	16.6	71.1	100.0	4.3	36.4	97.9	100.0	2.0	11.4	62.2	99.9
100	25	2.0	2.2	2.9	4.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
100	50	3.3	5.3	12.6	27.0	0.1	0.1	0.2	0.9	0.2	0.2	0.5	1.2
100	100	4.6	12.0	45.5	90.6	1.2	7.5	58.0	99.1	1.2	3.1	17.9	60.8
100	250	3.7	27.8	95.5	100.0	4.5	59.0	100.0	100.0	2.3	19.8	92.5	100.0
250	25	0.5	1.0	1.5	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
250	50	3.5	6.5	19.4	54.3	0.0	0.0	0.0	0.8	0.1	0.1	0.4	1.2
250	100	3.3	22.1	82.2	100.0	0.5	17.9	94.2	100.0	0.9	7.7	49.1	98.7
250	250	5.5	60.4	100.0	100.0	5.6	97.3	100.0	100.0	2.1	56.0	100.0	100.0

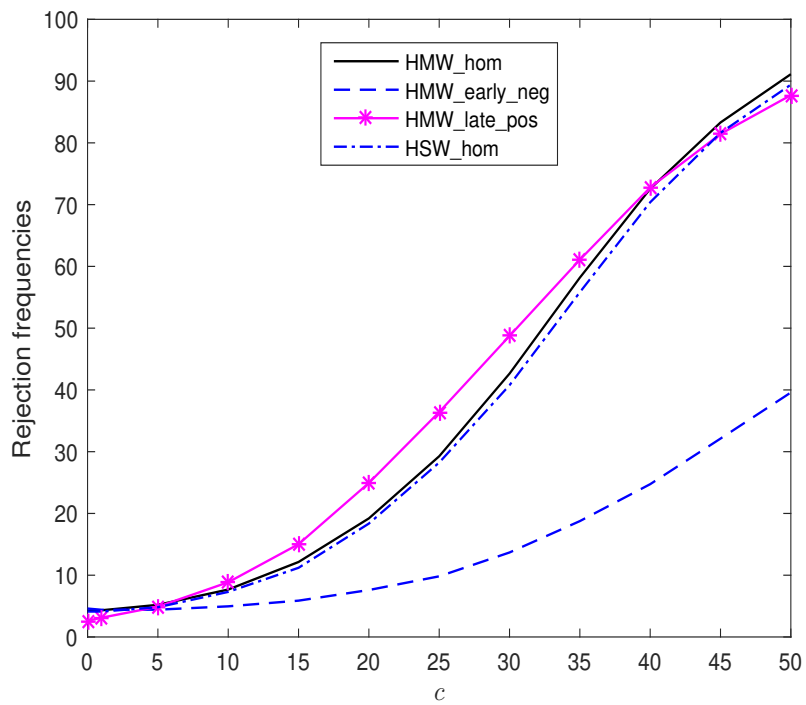
Notes: α denotes the share of stationary units in the sample, with the autoregressive coefficient of the stationary series generated as $\rho = (\rho_1, \dots, \rho_{N_1}), \rho_i \sim \text{iid } U(0.85, 0.95)$, where $N_1 = \alpha N$. For further notes, see Table 2.

Table 4: Is energy use per capita trend or difference stationary?

Period	Energy use p.c.						Ln (energy use p.c.)					
	y			Δy			y			Δy		
	$\hat{\tau}$	HSW	DH	$\hat{\tau}$	HSW	DH	$\hat{\tau}$	HSW	DH	$\hat{\tau}$	HSW	DH
<i>Full period</i>												
1960-2014	0.71	0.55	1.36	-2.79	-3.18	-2.56	1.56	1.24	1.37	-2.86	-2.88	-2.64
<i>50 years window</i>												
1960-2009	0.46	0.43	1.24	-2.65	-2.82	-2.17	0.86	0.91	1.39	-2.64	-2.40	-2.13
1961-2010	-0.18	-0.01	0.93	-2.82	-2.81	-2.21	0.77	0.84	1.34	-2.81	-2.59	-2.02
1962-2011	-0.18	0.00	0.40	-2.71	-2.81	-2.32	0.73	0.83	1.04	-2.69	-2.56	-2.01
1963-2012	-0.29	-0.10	0.18	-2.69	-2.85	-2.57	0.28	0.55	0.64	-2.76	-2.66	-2.17
1964-2013	-0.57	-0.36	-0.31	-2.67	-3.25	-2.61	0.04	0.39	-0.33	-2.84	-2.91	-2.74
1965-2014	-0.27	-0.14	-0.66	-2.70	-3.11	-2.18	0.19	0.48	-0.85	-2.79	-2.83	-2.51

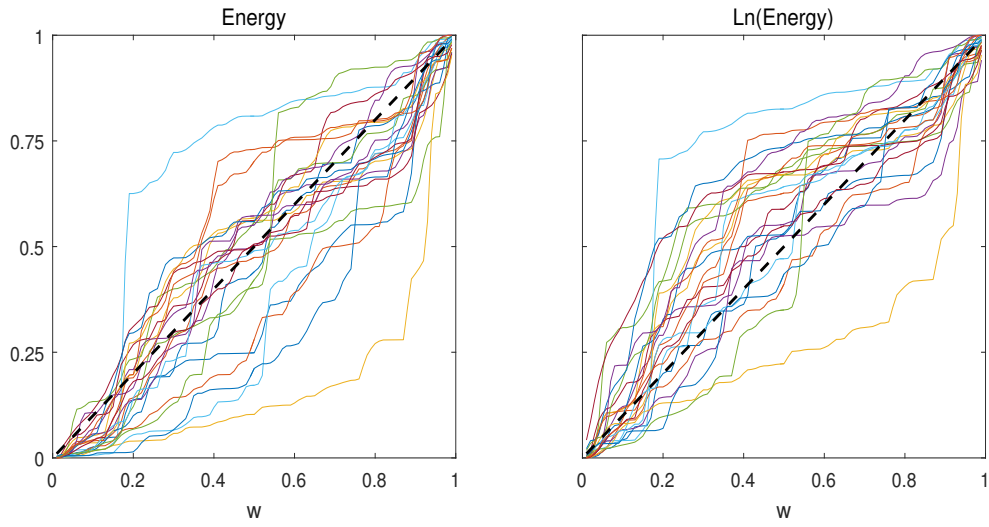
Notes: Reported numbers are estimates of the panel unit root tests $\hat{\tau}$, t_{HSW} and t_{DH} . Testing is performed on data that is first prewhitened and then recursively detrended. The lag order used for prewhitening is selected based on the AIC criterion, with the maximum lag length set to two. Our results remain qualitatively identical if we set a uniform lag length of one or two for all panel units. ‘Ln’ denotes the natural logarithmic transformation. Bold entries represent cases in which the panel unit root null hypothesis is rejected with 5% significance.

Figure 1: Local power of $\hat{\tau}$ and t_{HSW} as a function of c .



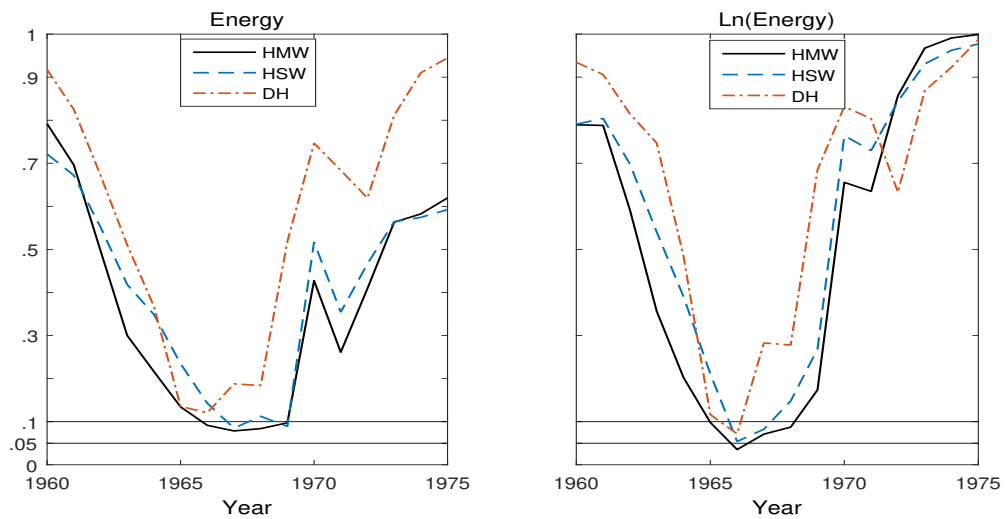
Notes: The curves depict local power of $\hat{\tau}$ (HMW) as a function of c under homoskedasticity, early negative variance break and late positive variance break. For the sake of comparison, the figure also depicts the local power of t_{HSW} (HSW) under homoskedasticity. Data are generated according to DGP1 and SAR(1) residuals. Regarding the selection of $\phi = 1 - c/(T\sqrt{N})$ we proceed similar to Westerlund and Breitung (2013) and draw $c \sim U(0, 2c)$. Rejection frequencies are computed based on 5,000 replications and a 5% critical value of 1.64.

Figure 2: Estimated variance profiles



Notes: 'Ln' denotes the natural logarithmic transformation.

Figure 3: Panel unit root testing over 40-years windows



Notes: The figures depict p -values from the panel unit root tests $\hat{\tau}$ (HMW), t_{HSW} and t_{DH} . 'Year' represents the year at which the 40-years sample period begins. For further notes, see Table 4.